

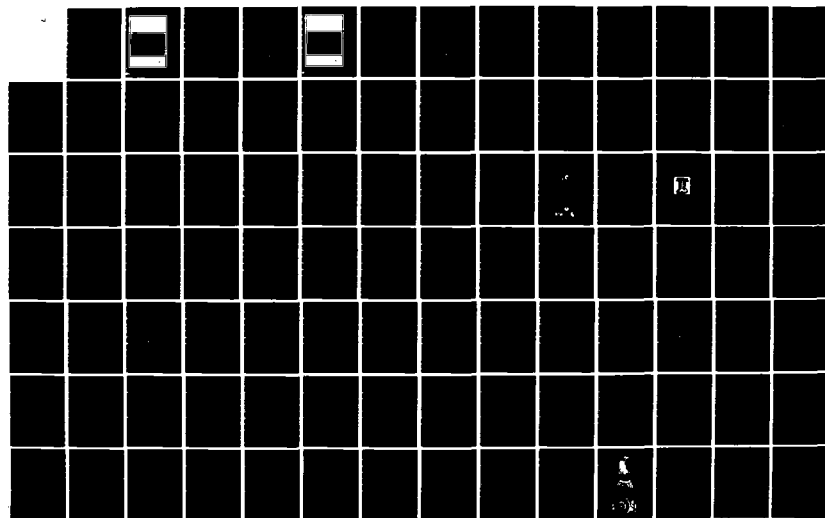
AD-A141 969

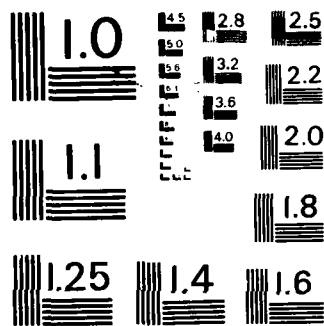
CONFERENCE PROCEEDINGS ON GUIDANCE AND CONTROL
TECHNIQUES FOR ADVANCED SP. (U) ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY... JAN 84
AGARD-CP-350 F/G 22/3

1/5

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A 141 969

①

AGARD-CP-350

AGARD-CP-350

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No.350

Guidance and Control Techniques for Advanced Space Vehicles

This document has been approved
for public release and sale; its
distribution is unlimited.

DTIC
MAY 1 0 1984
A

NORTH ATLANTIC TREATY ORGANIZATION



DISTRIBUTION AND AVAILABILITY
ON BACK COVER

84 05 10 046

DTIC FILE COPY

COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

(TITLE): Conference Proceedings on Guidance and Control Techniques for Advanced
Space Vehicles (37th) Held at Florence, Italy on 27-30 September 1983.

(SOURCE): Advisory Group for Aerospace Research & Development (AGARD)
North Atlantic Treaty Organization, 7 rue Ancelle, 92200 Neuilly sur Seine, France

TO ORDER THE COMPLETE COMPILATION REPORT USE AD-A141 969 .

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDINGS, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

AD#:	TITLE:
AD-P003 380	Controls Technology the Key to Future Space Systems
AD-P003 381	Spacecraft Control Research at Nasa
AD-P003 382	Image Dissector Tube Star Trackers Performance: Advantages and Drawbacks as Compared to those of Charge Transfer Device Star Trackers
AD-P003 383	The Fault Tolerant Spaceborne Computer
AD-P003 384	A Double Gimballed Momentum Wheel for Precision Three-Axis Attitude Control
AD-P003 385	Current Work on High-Grade Space Gyros at Ferranti in the UK
AD-P003 386	Overcoming Unobservability in Three-Axis Stabilisation of Satellites
AD-P003 387	Multifunction Spacecraft Attitude Estimation and Navigation System
AD-P003 388	A Filtered Attitude Determination System for Spacecraft Measurement and Control
AD-P003 389	Inertial-Optical Attitude Determination and Model Following Control of Manoeuvring Spacecraft
AD-P003 390	Docking Mechanism for In-Orbit Assembly and Spacecraft Servicing
AD-P003 391	The Influence of Sensor and Thruster Imperfections on the Attitude and Position of a Spacecraft Performing a Rendezvous
AD-P003 392	Control Aspects as Elaborated in Space Rendezvous Simulations
AD-P003 393	An Approach to Attitude Control Maneuvers which Helps Maintain Orbit inclination
AD-P003 394	Characterization of Uncertainty for Large Space Structure Control Problems
AD-P003 395	Attitude Control of Large Flexible Spacecraft
AD-P003 396	Low Authority Control Of Flexible Spacecraft Via Numerical Optimization
AD-P003 397	A Decentralized Active Control System for a Large Flexible Structure in Space

COMPONENT PART NOTICE (CON'T)

AD#:

TITLE:

AD-P003 398	Design and Robustness Analysis of Reduced Order Controllers for Large Flexible Space Vehicles
AD-P003 399	The Design, Simulation and Developmental Testing of the Space Data Bus System
AD-P003 400	Verification and Validation Techniques Applied to the Reconfigurable Block 5D Satellite Software

SELECTED
JUN 15 1984
A

Accession	7
Serial	
Index	
Classification	
Remarks	
Comments	
For	

This document has been approved
for public release and sale; its
distribution is unlimited.

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD CONFERENCE PROCEEDINGS No.350

Guidance and Control Techniques for Advanced Space Vehicles

This document has been approved
for public release and sale; its
distribution is unlimited.

DTIC
MAY 10 1984

NORTH ATLANTIC TREATY ORGANIZATION



DISTRIBUTION AND AVAILABILITY
ON BACK COVER

84 05 10 046

UNC FILE COPY

COMPONENT PART NOTICE

THIS PAPER IS A COMPONENT PART OF THE FOLLOWING COMPILATION REPORT:

(TITLE): Conference Proceedings on Guidance and Control Techniques for Advanced
Space Vehicles (37th) Held at Florence, Italy on 27-30 September 1983.

(SOURCE): Advisory Group for Aerospace Research & Development (AGARD)
North Atlantic Treaty Organization, 7 rue Ancelle, 92200 Neuilly-sur-Seine, France

TO ORDER THE COMPLETE COMPILATION REPORT USE AD-A141 969.

THE COMPONENT PART IS PROVIDED HERE TO ALLOW USERS ACCESS TO INDIVIDUALLY AUTHORED SECTIONS OF PROCEEDINGS, ANNALS, SYMPOSIA, ETC. HOWEVER, THE COMPONENT SHOULD BE CONSIDERED WITHIN THE CONTEXT OF THE OVERALL COMPILATION REPORT AND NOT AS A STAND-ALONE TECHNICAL REPORT.

THE FOLLOWING COMPONENT PART NUMBERS COMPRISE THE COMPILATION REPORT:

AD#:	TITLE:
AD-P003 380	Controls Technology the Key to Future Space Systems
AD-P003 381	Spacecraft Control Research at Nasa
AD-P003 382	Image Dissector Tube Star Trackers Performance: Advantages and Drawbacks as Compared to those of Charge Transfer Device Star Trackers
AD-P003 383	The Fault Tolerant Spaceborne Computer
AD-P003 384	A Double Gimballed Momentum Wheel for Precision Three-Axis Attitude Control
AD-P003 385	Current Work on High-Grade Space Gyros at Ferranti in the UK
AD-P003 386	Overcoming Unobservability in Three-Axis Stabilisation of Satellites
AD-P003 387	Multifunction Spacecraft Attitude Estimation and Navigation System
AD-P003 388	A Filtered Attitude Determination System for Spacecraft Measurement and Control
AD-P003 389	Inertial-Optical Attitude Determination and Model Following Control of Manoeuvring Spacecraft
AD-P003 390	Docking Mechanism for In-Orbit Assembly and Spacecraft Servicing
AD-P003 391	The Influence of Sensor and Thruster Imperfections on the Attitude and Position of a Spacecraft Performing a Rendezvous
AD-P003 392	Control Aspects as Elaborated in Space Rendezvous Simulations
AD-P003 393	An Approach to Attitude Control Maneuvers which Helps Maintain Orbit inclination
AD-P003 394	Characterization of Uncertainty for Large Space Structure Control Problems
AD-P003 395	Attitude Control of Large Flexible Spacecraft
AD-P003 396	Low Authority Control Of Flexible Spacecraft Via Numerical Optimization
AD-P003 397	A Decentralized Active Control System for a Large Flexible Structure in Space

COMPONENT PART NOTICE (CON'T)

AD#:

TITLE:

AD-P003 398	Design and Robustness Analysis of Reduced Order Controllers for Large Flexible Space Vehicles
AD-P003 399	The Design, Simulation and Developmental Testing of the Space Data Bus System
AD-P003 400	Verification and Validation Techniques Applied to the Reconfigurable Block 5D Satellite Software

DECLASSIFIED
JUN 15 1984
A

This document has been approved
for public release and sale; its
distribution is unlimited.

NORTH ATLANTIC TREATY ORGANIZATION
ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT
(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Conference Proceedings No.350
GUIDANCE AND CONTROL TECHNIQUES
FOR ADVANCED SPACE VEHICLES



THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
 - Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
 - Improving the co-operation among member nations in aerospace research and development;
 - Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
 - Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly
directly from material supplied by AGARD or the authors.

Published January 1984

Copyright © AGARD 1984
All Rights Reserved

ISBN 92-835-0349-X



*Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ*

PREFACE

Military applications of space for navigation, communication and intelligence impose increasing requirements on spacecraft capacity, orbit control and pointing accuracy. These requirements are to be met for a long active lifetime, with a high survivability and at low cost. As the size of the spacecraft tends to increase with capacity, in the near future large complex space systems will be needed, that require in-orbit assembly and which will lead to mechanical flexibility.

To meet the requirements for future spacecraft the performance of existing components, such as actuators and sensors, is improved or new concepts are developed. In particular the use of microprocessors and other data distribution systems permits multifunctional use of various sensors or information sources to produce effective, survivable systems at low cost. Increasing on-board computing capacity enables the use of sophisticated software for effective complex spacecraft control. A unique aspect of large spacecraft is the control of the structural configuration in order to achieve a specific pointing accuracy.

* * *

Les applications spatiales à des fins militaires dans les domaines de la navigation, des communications et des renseignements se traduisent par des exigences croissantes au plan de la capacité des vaisseaux spatiaux, du contrôle de leur orbite et de la précision du pointage. Il faut répondre à ces exigences en tenant compte de divers facteurs: longue durée de vie active, charges élevées de survie et prix de revient modique. La taille des vaisseaux spatiaux tend à s'accroître avec leur capacité; par conséquent, dans un proche avenir, on utilisera des systèmes spatiaux complexes et de grandes dimensions qui nécessiteront un assemblage en orbite et aboutiront à une souplesse mécanique.

Pour satisfaire aux impératifs des futurs vaisseaux spatiaux, on améliore actuellement les performances des composants existants, tels que servo-commandes et capteurs, ou l'on développe de nouveaux concepts. Le recours, en particulier, aux microprocesseurs et autres systèmes de distribution de données permet une utilisation multifonctions des divers capteurs ou des diverses sources d'information, en vue de la production, pour un coût modique, de systèmes efficaces et dotés de possibilités de survie. L'accroissement de la capacité de calcul embarquée permet l'utilisation d'un logiciel sophistiqué pour un contrôle efficace et complexe du vaisseau spatial. L'un des aspects uniques des vaisseaux de grandes dimensions est le contrôle de la configuration structurale dans le but de réaliser une précision de pointage spécifique.

GUIDANCE AND CONTROL PANEL OFFICERS

Chairman: Mr R.S.Vaughn
Director, Office of Research
and Technology
Naval Sea Systems Command
Washington DC 20362
United States

Deputy Chairman: Dr Ing. R.C.Onken
DFVLR, Institut für
Flugführung
Postfach 32 67
D-3300 Braunschweig
Germany

TECHNICAL PROGRAMME COMMITTEE

Chairman: Ir. P. Ph van den Broek, NE

Members: ICA D.Pichoud, FR
Dr R.C.Onken, GE
Ir. P.Kant, NE
Dr J.K.Sarnecki, UK
Prof. W.M.Hollister, US

HOST COORDINATOR

Major Castrucci
Scuola di Guerra Aerea
Istituto Militare Aeronautica
dell Cascine
Viale dell Aeronautica
Firenze, Italy

PANEL EXECUTIVE

Mr B.M.Heliot
AGARD NATO
7 rue Ancelle
F-92200 Neuilly sur Seine
France

(From USA and Canada only):
The GCP Executive
AGARD NATO
APO New York 09777

ACKNOWLEDGEMENTS/REMERCIEMENTS

The Panel wishes to express its thanks to the Italian National Delegates to AGARD for the invitation to hold this meeting in Florence and for the facilities and personnel which made the Meeting possible.

Le Panel tient à remercier les Délégués Nationaux de l'Italie près l'AGARD de leur invitation à tenir cette réunion à Florence, ainsi que pour les installations et le personnel mis à sa disposition.

CONTENTS

	Page
PREFACE	iii
PANEL OFFICERS AND PROGRAMME COMMITTEE	iv
TECHNICAL EVALUATION REPORT by R.W.F.Bain	vii
	Reference
KEYNOTE PAPER: CONTROLS TECHNOLOGY - THE KEY TO FUTURE SPACE SYSTEMS by R.J.Herzberg	K
SPACECRAFT CONTROL RESEARCH AT NASA by J.B.Dahlgren and L.W.Taylor	1
ROLE DE L'INTELLIGENCE DE BORD ET DE L'ENTRETIEN EN ORBITE DANS L'ACCROISSEMENT DE LA DUREE DE VIE DES SATELLITES par C.Cougnat et J.P.Sotta	2
PERFORMANCES DES VISEURS D'ETOILE A TUBE DISSECTEUR D'IMAGE: LEURS AVANTAGES ET INCONVENIENTS COMPARES AUX VISEURS D'ETOILE A DETECTEUR SOLIDE IMAGE DISSECTOR TUBE STAR TRACKERS PERFORMANCE: ADVANTAGES AND DRAWBACKS AS COMPARED TO THOSE OF CHARGE TRANSFER DEVICE STAR TRACKERS par H.Bourcier et M.Vite	3
REFERENCES GYROSCOPIQUES POUR APPLICATIONS SPATIALES* par J.Resseguier et G.Monier	4
THE FAULT TOLERANT SPACEBORNE COMPUTER by D.R.Bates, J.E.Schell and G.R.Couranz	5
A DOUBLE GIMBALLED MOMENTUM WHEEL FOR PRECISION THREE-AXIS ATTITUDE CONTROL by W.Auer	6
CURRENT WORK ON HIGH-GRADE SPACE GYROS AT FERRANTI IN THE UK by J.Wellburn	7
OVERCOMING UNOBSERVABILITY IN THREE-AXIS STABILISATION OF SATELLITES by L.Passeron and C.Bozzo	8
MULTIFUNCTION SPACECRAFT ATTITUDE ESTIMATION AND NAVIGATION SYSTEM by J.C.Amieux and B.Claudinon	9
A FILTERED ATTITUDE DETERMINATION SYSTEM FOR SPACECRAFT MEASUREMENT AND CONTROL by M.Noton	10
INERTIAL-OPTICAL ATTITUDE DETERMINATION AND MODEL FOLLOWING CONTROL OF MANOEUVRING SPACECRAFT by T.Zwartbol and A.P.Terpstra	11
DOCKING MECHANISM FOR IN-ORBIT ASSEMBLY AND SPACECRAFT SERVICING by F.Gampe	12

THE INFLUENCE OF SENSOR AND THRUSTER IMPERFECTIONS ON THE ATTITUDE AND POSITION OF A SPACECRAFT PERFORMING A RENDEZVOUS by R. Ancillotti and C. Cassi	13
CONTROL ASPECTS AS ELABORATED IN SPACE RENDEZVOUS SIMULATIONS by P. Natenbruk and D. Rangnitt	14
OPTIMISATION DES MANOEUVRES D'UN SATELLITE D'OBSERVATION EN VUE D'UN SURVOL RAPIDE D'UN POINT DU GLOBE* par A.M. Mainguy, C. Aumasson et S. Laffy	15
AN APPROACH TO ATTITUDE CONTROL MANEUVERS WHICH HELPS MAINTAIN ORBIT INCLINATION by W.M. Bowles and M.R. Altobelli	16
CHARACTERIZATION OF UNCERTAINTY FOR LARGE SPACE STRUCTURE CONTROL PROBLEMS by J.C. Doyle and J.E. Wall	17
ATTITUDE CONTROL OF LARGE FLEXIBLE SPACECRAFT by G. Thieme	18
LOW AUTHORITY CONTROL OF FLEXIBLE SPACECRAFT VIA NUMERICAL OPTIMIZATION by G. Schulz	19
A DECENTRALIZED ACTIVE CONTROL SYSTEM FOR A LARGE FLEXIBLE STRUCTURE IN SPACE by A. Danesi	20
DESIGN AND ROBUSTNESS ANALYSIS OF REDUCED ORDER CONTROLLERS FOR LARGE FLEXIBLE SPACE VEHICLES by S.S. Banda, D.B. Ridgely, H. Yeh and D.V. Palmer	21
THE DESIGN, SIMULATION AND DEVELOPMENTAL TESTING OF THE SPACE DATA BUS SYSTEM by A.J. Shapiro	22
LA CERTIFICATION DES LOGICIELS EMBARQUES: UNE APPROCHE GLOBALE ET PROGRESSIVE par G. Bracon	23
VERIFICATION AND VALIDATION TECHNIQUES APPLIED TO THE RECONFIGURABLE BLOCK 5D SATELLITE SOFTWARE by W.C. Diniak	24
MISSION CONTROL SYSTEM OF THE INFRARED ASTRONOMY SATELLITES (IRAS)† by R.C. van Holtz	25

* Abstract only. The full text is issued in Classified Publication CP 350 (Supplement).

† Not available at time of printing

TECHNICAL EVALUATION REPORT

by

R.W.Bain
Self-employed Consultant
32A Wellington Avenue
Fleet, Hampshire
UK

1. INTRODUCTION

1.1 The 37th symposium of the AGARD Guidance and Control Panel was held at the Italian Air Force's Air War College, Florence, 27-30 September 1983. For the first time the symposium dealt with spacecraft problems, the topic being guidance and control techniques for advanced space vehicles.

1.2 The subject has, of course, always been an important one in the continuing development of space systems for both military and civil purposes, the major problems having been to ensure a long active life for the spacecraft at a cost that was affordable while responding to mission requirements which have become more and more demanding over the years. In these respects military and civil systems are little different and it is perhaps not surprising that the symposium papers contained little of a specifically military nature even though the meeting was being held under military auspices. Another factor here, of course, is that the European members of NATO have so far been little involved in the development of military space systems.

1.3 Two things, perhaps, make the topic of particular significance at the present time -- first, the increasing trend to ever larger spacecraft with a multiplicity of appendages (large solar arrays, large antennas and so on) and second the impending deployment of large space stations assembled piecemeal in orbit, with the attendant need to develop rendezvous and docking techniques. Large structures, with their mechanical flexibility, present particular problems to the control engineer and control/structure interaction (CSI) is a driving force in many current programmes and figured largely in the papers presented. At the same time the massive leaps forward that have taken place in electronics technology, and which are continuing, have made it possible to contemplate the use, at reasonable mass, volume, power level and cost, of intelligent on-board systems with sophisticated and adaptable software which can ensure that effective guidance, control and stability is maintained. It was the objective of the symposium to examine such general issues as these.

1.4 The programme listed 25 papers in addition to the Keynote Address. Of these, three were not presented, in one case a substitute paper being given by the Session Chairman, and one additional paper was inserted into the programme. There were six technical sessions.

2. OPENING CEREMONY AND KEYNOTE ADDRESS

2.1 Opening Ceremony

2.1.1 Following initial remarks by Mr R.S.Vaughn, Chairman of the Guidance and Control Panel, the symposium was formally opened by Major General G.A.M.Marconi, the Italian National Delegate to AGARD. He stressed the importance generally of AGARD as a collaborative method of producing the technology required for military operations and underlined the importance of the symposium topic as a key factor in the design of spacecraft. General D.A.L.Meloni, Commandant of the Air War College, who followed, also noted the importance of guidance and control techniques in spacecraft design and the need to develop low cost and effective solutions to the many problems. Mr R.K.Geiger, Director of AGARD, presented the thanks of AGARD to the Italian Air Force and to the Air War College for their efforts in staging the meeting and for their hospitality. He pointed to the many opportunities available for collaborative research in the general field of space systems for military purposes, noting that the capabilities of these systems were such as to provide for the needs of more than one group. Symposia such as the present one were valuable in exposing new lines of thought: the papers presented were important, but the opportunities available for discussion and for the making of contacts with other experts were probably more so.

2.1.2 The meeting was then handed over to the Technical Chairman, Mr P. Ph. van den Broek, from the Department of Aerospace Engineering at the Delft University of Technology, who, before introducing the keynote address speaker, commented on this being the first meeting of the Guidance and Control Panel on space-related topics. It was a timely

meeting, given the increasing importance of space systems in military operations, and, of the many facets of space technology, guidance and control techniques were an important element.

2.2 The Keynote Address

2.2.1 In the absence of Mr Dell P. Williams III (Lockheed Missiles and Space Co. Inc.) who had been the nominated speaker, the keynote address was given by Mr R.J. Herzberg (of the same organisation). Mr Herzberg introduced himself as a generalist in space technology – he was responsible for space technology planning at Lockheeds – and not as an expert in guidance and control techniques, which, nevertheless, he saw as one of the critical areas. Space was now twenty-five years old and a first and fundamental question that had to be answered at the present time was whether further advances and extensions were required. NASA, the US Air Force and the US aerospace firms were all concerned in these matters but looking ahead to what was necessary was difficult since evidently for financial reasons all the possible programmes could not go ahead. He cited a number of the more important future programmes and identified the two significant control technology issues as

(a) the need for the active suppression of flexible modes in large satellites and space stations of the future

and

(b) the need for increasing on-board autonomy.

Of these, the first was the more difficult and the whole business of control/structures interaction was currently a strong research area. A NASA committee had concluded that a number of future missions would suffer serious constraints if new techniques were not developed. As to autonomy, he felt that on-board navigation, station keeping and control was mainly needed by military systems, reducing as they did the ground stations' costs and adding to the protection of spacecraft in orbit.

2.2.2 He then went on to point out the dilemma of the programme planner – the problem generally of deciding which technologies were flight ready, providing minimum risk at reasonable cost – and noted the conflicting views of the programme managers (pessimists who preferred to rely on well proven techniques) and the research engineers and scientists (optimists with confidence in the new approaches).

2.2.3 Returning to the problems of flexible body control, he emphasised the need for good modelling but also the necessity to be able to reduce the complexity of the model, retaining only the essential system parameters. Currently, the three areas of expertise – control technology, computer technology and structural design – were separate disciplines and he foresaw a new type of engineer emerging by the end of the decade with the ability to apply all three.

3. FIRST TECHNICAL SESSION

3.1 The session was chaired by Mr van den Broek and was entitled 'Review of Mission Requirements and Technology Issues'.

3.2 The first paper, 'Spacecraft Control Research at NASA' (Paper 1) was presented by Mr J.B. Dahlgren and provided an outline of what were considered in NASA to be the basic problems in the control and guidance of future space vehicles. These, in general, related to the need to control spacecraft which would be both large and flexible. As far as theory was concerned, a number of deficiencies had already been identified and these were being investigated. There was also the problem of the large structure where multiple sensors and actuators were required to maintain the configuration. In the field of adaptive control, techniques were being investigated for application to structures where unknown or changing parameters might require modifications to the control system design. The space station was a good example of this. Assembled step by step in orbit, its mass would change, its centre of gravity move, there would be varying external torques and vibration modes would alter throughout its life. The conventional approach – rigid body with flexible appendages as applied to Skylab and the Shuttle – might suffice for the initial configuration but as the station grew the structural dynamics would make an increasing degree of adaptive control necessary. The required theoretical base thus had to be established. As far as spacecraft technology was concerned the current programme dealt with large antennas and with the requirements of submillimetre and infra-red astronomy satellites. However it was clear that there were future requirements which called for several orders better surface shape and pointing accuracy. There was also the problem of the large deployable reflector for astronomical observation – to satisfy the requirements it would be necessary to measure the position and orientation of each section of the antenna and to control these quantities with very high precision. There was need therefore in this and other applications for a very accurate shape determination system and for control actuators. Returning to the control of flexible structures, the limitations of ground experiment were noted and hence the need arose for flight experiments – a series of such experiments in the late 80s and early 90s were currently being planned together with a set of ground tests for comparison.

3.3. The presentation and even more so the paper itself which became available later in the week of the symposium covered a great deal of ground and illustrated the major problems which will have to be overcome in controlling the large flexible spacecraft of the future. There was a brief discussion period during which the problem of control in the presence of propellant sloshing was identified as another area of difficulty.

3.4 The final paper (Paper 2) of the session, by Cougnet and Sotta, dealt with the role of on-board intelligence and in-orbit servicing as means of increasing the lifetime of satellites. After listing the normal limitations to satellite life and a number of possible remedies, the paper went on to consider in very general terms the two topics selected.

3.5 As to the first, the high reliability of modern micro-computers having been noted, their uses in a number of on-board tasks were examined – power control and distribution, thermal control, propellant management, the monitoring of sensor performance, the adaptation of control laws as the external environment changes, data management and so on. Potential gains were listed as well as the problems – the need, for example, to account for the reliability of the micro-computer itself and of the cost of the on-board intelligence.

3.6 The authors then went on to discuss the use of on-orbit servicing in such matters as re-fuelling, repair and replacement of subsystems and the modification of the payload as mission needs changed with time. They pointed to this as an attractive and flexible solution but also noted the inherent difficulties and the need to develop a wide range of new techniques and systems – rendezvous and docking, robotics, the service vehicle itself and so on – as well as to take account of the implications of servicing on the design of the satellite itself. The question of man's presence was considered briefly.

3.7 There was very little time for discussion and the only question brought out that the authors were thinking in terms of satellite life times in the range 15 to 20 years. The paper was, in fact, a very general one but at least it conveniently listed the problems to be solved and the potential and difficulties of the two methods of solution postulated.

4. SECOND TECHNICAL SESSION

4.1 This was entitled 'Control Components' and dealt largely with existing technology. One of the papers advertised on the programme (*Peacekeeper Missile Thrust Vector Control System*) was not given but an additional paper by staff of Ferranti, UK on their current work on space gyros was presented, immediately following a similar paper by staff of SAGEM, France.

4.2 The first paper (Paper 3) by Bourcier and Vite of SODERN, France, considered the relative merits, as used in star sensors, of image detector tubes and charge coupled devices, taking as an example of the former the family of sensors developed for use in Exosat and the Spacelab Instrument Pointing System (IPS). The particular sensor developed for the IPS was described in some detail and its main characteristics as determined by experiment were given. The paper compared the performance of the image dissector tube with that of a particular RCA charge-coupled device, examining generally the three parameters – angular resolution, acquisition time, and field to accuracy ratio – as well as their convenience in use. As usual in such comparisons there were pros and cons either way and no clear preference emerged.

4.3 There was unfortunately no time for questions or discussion.

4.4 Two papers followed on the subject of space gyros. The first of these (Paper 4) by Resseguier and Monier of SAGEM, described in some detail the characteristics of the SAGEM floating gyros FO6 and GYPSE which have already been used or are intended for use in a number of European space programmes and also those of the dry tuned gyros, GILDAS 1 and 3, of which the space use had been the subject of recent study. A broad conclusion of a comparison of the two types was that the dry tuned gyros had the greater potential provided that the production techniques already established for other applications did not have to be modified to any great extent. In passing it was noted that laser gyros were not greatly favoured. The paper concluded with a study of the optimum geometry of a gyro sensor cluster, together with some mention of the value of coupling star sensors with the gyro package.

4.5 The discussion that followed raised one or two questions of fact and also underlined the SAGEM view that the greatest potential in the immediate future lay with the dry tuned gyro. It was felt that perhaps 20 years hence the laser gyro could come into use. The paper itself was labelled 'NATO Confidential'. It is a little difficult to see what the military security implications were and there appeared to be little in the paper that would not have been published in the standard Company brochure.

4.6 The paper that followed (Paper 7), an addition to the published programme, was presented by Mr Wellburn and discussed current work at Ferranti, UK, on high grade space gyros. Although most US companies had switched from floated gyros to dry gyros, Ferranti's view was that the former were worthy of continued development and improvement. The Type 125 had been much used and successfully so in a number of European spacecraft and launch vehicle programmes and laboratory tests of the lifetime of such ball-bearing gyros confirmed their suitability for long space missions. The Type 126, gas bearing gyro, on which Ferranti had also worked, while suitable for many aircraft and launcher applications, was still not capable of achieving the performance of the best floated gyros.

4.7 In a brief question time, the Ferranti view that the technology of fibre optic gyros had still a long way to go was brought out. It was generally interesting in relation to both the SAGEM and Ferranti presentations that there was strong belief in the reliability of ball bearings and that such mechanical devices had a long life ahead in space applications.

4.8 The next paper by Bates, Schell and Couranz (Paper 5) described a fault-tolerant spaceborne computer developed by the Raytheon Company for the USAF and intended to meet the on-board computational and survivability needs of space missions in the 1983 to 1993 timeframe. It had not so far been used for a space application but the techniques employed had been utilised on computers used by the US Army. The paper was presented with the aid of a useful film. The computer was described as having been designed to support a variety of space missions with a 95% probability of surviving without performance degradation for 5 to 7 years. It was segregated into a number of functional elements, each of which was described in some detail in the paper, with each element protected by at least one spare. In addition a number of these elements was provided with internal redundancy through the use of what were called Rippler switches these were a key new feature in the design. The capability of the computer for self-adjustment following internal faults or radiation events, its value for autonomous control, the consequent relaxation of ground station requirements and its ability to be reprogrammed simply from the ground were all stressed. It was strongly urged that this was not just a paper computer but that extensive testing had proved its worth.

4.9 In discussion the question of fault tolerant software was raised. The speaker's view was that full fault tolerance in software was extremely difficult to achieve and there was still no substitute for very careful design and checking of programs. In answer to another question on the effect of radiation, the speaker outlined the many-faceted approach the protection of memories using error correcting codes and the simple expedient of periodical memory checks, for example, as well as the technique of recovering from short power breakdowns the ROM could be used to establish the position at breakdown and lead to a resumption of computation. The key role of the rippler switches was protected by making them as simple as possible. In all this was an interesting and well presented paper, only lacking in the fact that no space use of the device had yet occurred.

4.10 The session ended with a presentation by Dr W.Auer of TELDIX (Paper 6) on a double gimballed momentum wheel for precision three-axis attitude control. The paper was well presented and contained a detailed technical description of the device, its general principles of operation and an outline of the system tests carried out by DFVLR. The most critical subassembly was identified as the ball bearing unit but here again, as in the case of the gyros, there was every confidence in its reliability and life-time in space applications. The main message was that here was a device which had been developed and qualified for space use and it was available. There were no questions.

4.11 The session was perhaps a little disappointing in that it confined itself to descriptions of currently existing equipment and provided little more than could be found in contractors brochures. Greater emphasis on possible future developments in the field of control components would, it is felt, have enhanced the worth of the session. Apart possibly from the paper on the fault-tolerant computer, which of course may have been new only to European ears, there was little of a thought-provoking nature.

5. THIRD TECHNICAL SESSION

5.1 This session was chaired by Dr A.J.Sarnecki of the Royal Aircraft Establishment, UK, and was entitled 'Estimation and Control'. The programme listed five papers; four only were presented, the fifth, by Turci et al. (on attitude control and stabilisation techniques for advanced communication spacecraft utilizing an on-board processor) having been apparently withdrawn.

5.2 The first paper (Paper 8) by Passeron and Bozzo considered the problem of overcoming unobservability in three-axis stabilization of satellites, taking as their model the well-known geostationary three-axis stabilised satellite using the pitch momentum concept. Thus the satellite comprised a central body, with rigid appendages and a sun pointing solar array, and a fixed momentum wheel along the pitch axis. The on-board sensors consisted of an infra-red sensor providing noisy measurements of pitch and roll and a sun sensor providing, once or twice per orbit, noisy measurements of yaw. The promise was that in future spacecraft much more accurate control in both yaw and roll would be required. The objective of the paper was to demonstrate how this could be accomplished, without an accurate yaw sensor, by including in the spacecraft model a representation of the expected roll and yaw torques. These were represented by truncated Fourier series, the fundamental frequency being that of the Earth's rotation. With that built into the design of the filter, it appeared from their simulation results that a marked improvement could be achieved in both roll and yaw angle control with only moderate fuel consumption. The analysis presented in the paper dealt only with the first term in the Fourier series but it was stated that the method could be readily adapted to take account of higher harmonics.

5.3 In answer to questions it was agreed that if the fundamental frequency was wrongly estimated, then the method would be in difficulty. It was not possible to extend the technique to cover pitch angle control (agreed not to be too much of a problem in any case) since the pitch torque had discontinuities although the question had been looked at in a recent response to an ESA study.

5.4 The next three papers all considered various aspects of optics-inertial systems for the estimation of spacecraft attitude. The first paper (Paper 9) by Amieux and Claudinon essentially summarised five years of research at Matra on the subject, during which time the general concept has been refined and improved. After introducing the general principles of the technique, listing the various sensors (infra-red Earth sensors, sun sensors, star scanners and trackers and radiofrequency sensors) which could be used in concert with the gyro package, the authors went on to consider three specific case studies. The first concerned a large geostationary communication satellite (work carried out for

Intelsat) where there was a need for a direct yaw measurement. The attitude measurement assembly comprised an infra-red Earth sensor, a three-component gyro pack and a digital sun sensor. Results of simulations demonstrated a successful design with yaw being controlled within 0.05° . The second example (work for CNES) related to an Earth observation satellite of the SPOT type (in low sun synchronous orbit) where an assembly comprising a star mapper and a six-gyro set had enabled attitude control to be maintained within ten arcsec. Finally the problem of the state estimation of a flexible spacecraft where the auxiliary sensors comprised an infra-red Earth sensor and a radio frequency sensor was currently under study. A number of satellite missions already using the concept or being designed to do so (Exosat, SPOT, Hipparcos) were noted. The authors concluded with outlines of the possible application of the technique to the two problems of autonomous station keeping and on-board relative trajectory estimation during a low Earth orbit rendezvous, both of them interesting problems for the future.

5.5. Dr M. Noton of British Aerospace Dynamics Group in his paper (Paper 10) which followed, gave a progress report on a UK collaborative project, FADS (filtered attitude determination system), where the measurement system was based on a gyro package and a star crossing detector which was expected to be lighter and less power consuming than the more usual star tracker. The project involved collaboration between BAeDG (system studies, microprocessor software and hardware), MSDS (star sensor development and system assembly, integration and test), and Ferranti (gyro development and testing) and the immediate aim was to demonstrate the system at breadboard level. An important feature was compatibility with ESA's modulator attitude control system (MACS). Later work would be aimed at a future system for flight application. Thus, for example, the computer chosen for the breadboard demonstration (Intel 8086) would be replaced by a flight qualified device. The paper discussed the star crossing detection design based on the Fairchild 1728 element CCD array, the choice of on-board processing for the star tables, results of recent tests by Ferranti on the Type 125 gyro as well as outlining the derivation of the filter algorithm, made more difficult in this case by the intermittent nature of the information provided by the star crossing detector. The system appeared to be suitable for either near-Earth or geostationary Earth pointing satellites providing a pointing accuracy of the order of two arcsec. Future developments of the system were expected to consider the use of star trackers or radio frequency sensors.

5.6 The final paper of the three (Paper 11) by Zwartbol and Terpstra of NLR, Netherlands, discussed similar work at NLR on inertial-optical attitude determination and the use of model following for the control of manoeuvring spacecraft. The control system considered in the paper consisted of a strap-down rate integrating gyro, a star slit sensor and a reaction wheel actuator supported by an on-board computer. The paper considered in turn the control law, based on a simplified spacecraft model, and algorithms for spacecraft state determination and for the estimation of gyro parameters. Software simulation had indicated the satisfactory performance of the algorithms, but of much more interest had been the fact that the control law and spacecraft state estimator had been tested in flight using the Infra-red Astronomy Satellite (IRAS). The algorithms had been loaded from the ground into the IRAS computer and control had been taken over from the existing system. Satisfactory performance had been demonstrated though there had been some discrepancies as compared with the ground simulation. These were in fact simply explained by a mismatch in the moment of inertia of the spacecraft model and by a delay between the sampling instant and the issue of a new command which had not been incorporated in the simulation. This was an interesting experiment, a real test of the system with a real satellite, and much more convincing than simulation where one is generally testing the system designed in relation to the model assumed for the design.

5.7 For the purpose of questions, the three papers were grouped together and a good discussion ensued, within the time available. This is an interesting subject and the general technique of mixed inertial and other sensors has already found application in a number of satellite programmes and will continue to do so. The papers formed a useful addition to knowledge of the subject.

6. FOURTH TECHNICAL SESSION

6.1 This, chaired by H. Radet, France, was entitled 'Orbital Manoeuvring'. Five papers were presented in a slightly different order to the published programme.

6.2 The first of these (Paper 14) by Natenbruk and Rangnitt considered some of the control problems of space rendezvous as brought out by digital simulations. The paper noted the increasing interest in the topic with the beginning of the operational use of both the Space Shuttle and Ariane and referred to three specific cases - the demonstration of a rendezvous and docking (RVD) capability in low Earth orbit using the Shuttle, the attachment of an apogee motor to a Comsat in geostationary transfer orbit using Ariane 3 and the assembly in geostationary orbit of a modular Comsat, using Ariane 4. The main business of the paper, however, was to describe the simulation programme which dealt with the final phase of the rendezvous, the early phases being assumed to be dealt with by ground station control, and an outline was given of the flexibility of the programme in dealing with, for example, different control laws (constant bearing courses or pursuit courses) and with the problem of 'the last few metres'. Two particular examples of different ways of dealing with the final approach problem were discussed. It was clear that further work still remained to be done to arrive at the best solutions to the problem but evidently the simulation programme which had been devised was an essential tool in further study.

6.3 The second paper (Paper 13) by Ancillotti and Cassi, dealt with the effect of sensor and thruster imperfections on the attitude and position of a spacecraft performing a rendezvous. The particular case examined was that of rendezvous

in geostationary orbit, the target being a rigid spacecraft controlled from the ground but equipped with laser radar instrumentation and the chaser a rigid three-axis stabilised spacecraft with an optical telescope, corner reflectors and a sun sensor for attitude measurement. The paper developed in turn the mathematical model of the system, the position and velocity control laws, the chaser attitude control laws and the chaser position and attitude measurement technique and then went on to examine the effect of sun sensor errors and thruster misalignment on the rendezvous manoeuvre. The sensor errors assumed appeared to have little impact, with position and velocity measurement errors being more significant. The need to match the thruster minimum impulse to the sensor errors was stressed.

6.4 The third paper (Paper 12) by Gampe, turned from the realm of theory to the practical design of a docking mechanism for in-orbit assembly and spacecraft servicing. The docking mechanism subsystem was of course recognised as only one of a large number of technology items and techniques which would be needed for future tasks such as spacecraft assembly, the maintenance, repair and refurbishment of satellites in orbit and so on but it was an essential element in the final rendezvous phase. The paper looked first at the general requirements in relation to different docking strategies and then considered in some detail the docking mechanism itself and the associated electronic system. The hope was for a first demonstration of docking during the ESA's Eureka 2 mission at the end of this decade.

6.5 In the discussion that followed the three papers a number of points were brought out:

- (a) the control laws studied were simple, not necessarily the best but were realistic,
- (b) minimising propellant consumption during the final phase was probably not important since it only represented a small fraction of the total needed for the whole rendezvous mission,
- (c) the simulation work done was at an early stage in Europe and there could be a harder look at the computation process before a proper design was implemented,
- (d) there might be a need to include, in the relative motion equations, some allowance for differential orbit changes between chaser and target (the two theoretical papers appeared to differ on this score) — for example drag effects in Eureka with its large solar array could be important,
- (e) an estimate of the computational load during final docking ought to be made,
- (f) a cruder docking mechanism capable of standing up to higher closing velocities might be a preferable solution.

The discussion illustrated very well the current strong European interest in the topic but equally that studies were at an early stage and many trade-offs still had to be performed. It was perhaps a little surprising to find no US input on this subject — possibly their work has already been well publicised.

6.5 The next paper (Paper 15) by Mainguy, Aumasson and Laffy looked at the problem of optimising the manoeuvres of a surveillance satellite aimed at achieving a quick repeat over-flight of a particular point on the Earth. Three methods of analysis of increasing complexity were presented, with different assumptions about thrust level and direction and about the possibility of linearising the equations of motion about the initial orbit.

6.6 In a later question period, the assumption made in the paper, that the final orbit was in the same plane as the initial, was discussed. Usually a surveillance satellite would be in a synchronous orbit. If the height of the orbit had to be changed then to retain sun synchronisation an inclination change would be needed. This had not been studied but it was thought that the third technique discussed could include this feature. The paper itself gave a good introduction to the topic.

6.7 It was a little surprising that the paper was classified NATO Confidential since there appeared to be little of a military sensitive nature exposed.

6.8 The session finished with a neat short paper (Paper 16) by Bowles on a method of performing attitude control manoeuvres for a spinning satellite in such a way as to help maintain the orbit inclination. A spin stabilised satellite is normally controlled in attitude by firing axial thrusters (parallel to the spin axis). The orbit plane for a geostationary satellite is also maintained close to zero inclination by thrusting axially as the satellite passes through the orbit's descending node. Thus if attitude control can be exercised close to the descending node, the thruster impulses will at the same time contribute to the correction of the orbit inclination. The paper described the procedure clearly and one numerical example given indicated that, with this method, out of 14.2 kg of propellant required by Intelsat VI for attitude manoeuvres over a ten-year life, 13.0 kg would contribute to inclination control — a useful bonus at no great difficulty.

6.9 A short discussion before the start of the next session clarified one or two matters of fact.

7. FIFTH TECHNICAL SESSION

7.1 This session was chaired by Dr T. Cunningham of Honeywell, USA, and under the title of 'Large Space Structures' comprised five papers of a highly theoretical nature. The paper by Schulz (Paper 19) was not given due to the author's illness but the gap was filled by Dr Cunningham himself.

7.2 The first paper (Paper 17) by Wall and Doyle dealt with the characteristics of uncertainty for large space structure control problems. A preprint of the paper was not available during the symposium due to its not having been cleared for publication. The authors discussed work carried out for USAFOSR where the objectives were, for a large space structure, to examine different ways of describing the uncertainty (in, for example, sensor and actuator characteristics and in structural properties), to compare methods of analysis and to propose a formulation for the required feedback control system, which would function adequately in this situation. The method of modelling the various uncertainties was described, analysis methods compared and the structured singular value introduced as a method of analysing performance in the presence of uncertainty.

7.3 The second paper (Paper 18) by Thieme entitled *Attitude Control of Large Flexible Spacecraft* described some initial results of a study carried out for ESA. Three models were used - a free beam, a flat plate with a rigid centre body and three coupled plates. None, of course, represented a real or planned satellite but they possessed the characteristics which were important in the design of attitude control systems. The work had reached the stage of having completed the analysis of requirements, the setting up of the dynamic models and their analysis, and the design of the controller. Later work would aim at evaluating performance and the derivation of an experiment and test programme.

7.4 To fill the gap left by Paper 19, Dr Cunningham gave a presentation on some work on the precision pointing of flexible space structures, the main problem being to achieve a specific mission performance in the face of unknown (or imperfectly known) spacecraft characteristics and of uncertain disturbances. Four classes of control problems were discussed - conventional attitude control taking account of flexibility with bending modes in the control passband, secondly, conventional with vibration isolation, thirdly, attitude control with isolation and vibration control and finally integrated pointing and vibration control. As a broad overview, the presentation gave a clear account of the problems and of potential solution methods.

7.5 Danesi followed with Paper 20 in which he described a decentralised active control system for a large flexible structure in space. The particular configuration studied consisted of a symmetrical structure with a central hub and arms carrying solar panels and large communications antennas. In the work done so far only the arms were assumed to be flexible. The paper described the structural model used, the method of analysis and the control system devised. Some simulation results showed the feasibility of the approach.

7.6 Finally, Paper 21 by Banda et al. discussed the design and robustness analysis of reduced order controllers for large flexible space vehicles. The speaker went over three topics - the background to reduced order control and the use of linear-quadratic-Gaussian methods in control design, stability robustness analysis where the use of singular value theory was noted, and performance analysis. It was pointed out that the several stability robustness tests did not imply each other and that while the passing of a test did not guarantee control stability for a given design, failure did not necessarily imply the opposite. Some simple examples, not representative of real physical systems, were discussed and analysed.

7.7 There was very little discussion of these papers, only one - Paper 18 - attracting a couple of questions from the Chairman. Such papers, however, with a high mathematical content are difficult to present and difficult for an audience to assimilate, especially without a preview of the paper. A proper study of the papers when available could well help in understanding the applicability of the theory to real practical problems.

8. SIXTH TECHNICAL SESSION

8.1 This session, chaired by P. Kant, Netherlands, turned from high theory to the very practical area of test and performance evaluation. There were four papers, a preprint of the last one not being available.

8.2 The first paper (Paper 22) by Shapiro discussed the design, simulation and development testing of the space shuttle data bus system. The paper was very clearly presented and the paper itself gives a good account of the various stages of development - the system design taking account of the vehicle needs, the detailed design of the system modules, each of which was described in turn, and the analysis and simulation of the complete unit. Future trends in such design were seen to be towards a mixture of wire and optical fibres to deal with the high data rates (10-20 Mbits per sec), essential in future aircraft, spacecraft and space stations. Much study of the future was under way by such bodies as the SAE AE-9B High Speed Data Bus Subcommittee and an examination was being made of various architecture, topology and protocol candidates. The paper itself lays down a good basis for the design, analysis and test of future systems.

8.3 Questioned on the criteria for going to fibre optics, the speaker noted that so far such technology was not at an operational stage - there was a lack of suitable components and much need for standardisation. The technology would be used, however, in the future for many high speed functions, such as on-board radar data processing.

8.4 The second paper (Paper 23) by Bracon discussed the problem of the validation and verification of on-board software. The purpose of validation and verification, of course, is to be sure that the software is capable of accomplishing every specified function with the required precision under the prescribed operational conditions. The particular example of the main software developed for the Mirage 2000 was discussed with reference to the successful procedures developed and future trends and needs were examined.

8.5 In discussion it was noted that mission duration was important in determining the level of test required and also that the use of a high level language did not eliminate the need for certification.

8.6 Paper 24 by Dimak on verification and validation applied to the USAF Defense Meteorological Satellite Program (DMSP) on-board software, though presented out of order and last, is best discussed now since it relates closely to the previous paper. The speaker began with a useful description of the DMSP system and of the satellite itself, together with a brief history of the spacecraft so far placed in orbit, before going on to the techniques used for verification and validation - the first he defined as ensuring that the computer programme had been developed in accordance with the software specification and the second as determining that the software would perform the necessary functions in the operational environment. Current verification procedures were discussed as well as validation methods using the multi-processor software test facility. The capability to modify or reconfigure the on-board software to cope with spacecraft subsystem failure or malfunction was noted and reference was made to a number of cases where V and V had revealed software faults which could have had catastrophic results in orbit. He was confident that at least in the DMSP context verification and validation of software had more than paid its way.

8.7 In a brief question period the use of the software simulation model was described as valuable when doubts arose about the behaviour of the in-orbit satellite. Clearly, the whole business of software test and verification will become more and more important in the future as more use is made of microcomputers for on-board tasks. Confidence in the hardware is perhaps more easily achieved than in the software and the two papers were very useful in underlining the existence of the problem and in describing the techniques and methods which have already been successfully applied.

8.8 The final paper on the programme (Paper 25) was presented by van Holtz and dealt with the mission control system of the infra-red astronomy satellite (IRAS). The speaker gave an outline of the satellite and its payload and of the ground operations required. A particular feature of the satellite was the ability to reprogram the on-board computer from the ground and this facility had been used at least once per month from launch. It had been possible by this means to overcome the problems caused by the anomalous behaviour of one of the on-board sensors early in the life of the satellite and as another example, the sensitivity of the infra-red telescope to Earth radiation had been controlled by means of on-board programming. Other capabilities - recovery from danger situations and the in-flight test of new control methods (described in Paper 11) were mentioned. It was noted in the question period that the new control technique once tested could readily be substituted for the old. A further example of the usefulness of the facility related to the previous Dutch astronomy satellite, ANS, where, the satellite having arrived in the wrong orbit, all the on-board software had had to be changed. The need, of course, to test the software by ground simulation before on-board use was stressed.

8.9 This was a useful and interesting session dealing with real solutions to practical problems in an area of space system design which will undoubtedly become more important in the future. Useful pointers to future work were provided by the various speakers.

9. ROUND TABLE DISCUSSION

9.1 A short discussion period followed the final technical session, the platform party comprising Miss Eveline Gottzem, Mr J.B. Dahlgren and Dr T. Cunningham, with Mr P. van den Broek as Chairman.

9.2 Each of the panel members made an opening statement and discussion centred around the points they made. There was the view that once again the gap between theory and its practical application had widened - this particularly exemplified by the papers of the fifth technical session. It was felt that earlier advances in control theory had been helpful, not so much in aiding system design as in assisting the formulation of the problem. There was evidently no universal recipe for solving control problems, no automated control system synthesiser, but if theory could help in establishing a proper statement of the problem (for example by devising mathematical ways of describing system uncertainties), then this was at least a step forward. The method of solution then did not matter greatly, though the point could be made that in formulating the problem (e.g. the performance criteria) the method of solution was dictated. There was almost bound to be a gap between theory and practice, with the practitioner taking time to see the relevance

the transfer of the state space methods of the early 60s had been slow but the techniques had been found to be useful and there was no reason to suppose that this would happen again with the applicable elements of the new theory

9.3 The importance of the subject of interactions between structure and control systems had been underlined by the number of papers on the topic presented at the symposium. There was need for further development of modelling and model reduction techniques and for cost effective ground and flight test methods. Of course, active disturbance control need not always be required and for example there would always be a possible trade-off between stiffening the structure and complicating the control system.

10. COMMENT ON SYMPOSIUM ARRANGEMENTS

10.1 The list of attendees issued at the symposium indicated that around 130 people had registered. The audience at the early sessions probably averaged about 75 with a tendency to a smaller number towards the end.

10.2 Twenty-four papers, in addition to the Keynote Address, were presented in six technical sessions and the majority of them were of a good standard. It was unfortunate that preprints of the papers were available at best just before presentations were made and an earlier availability certainly in the case of the more abstruse papers, might have led to livelier discussion periods which were generally short.

10.3 That the discussions were short was in a number of cases due to presenters over-running their allotted time – a point made by the Symposium Chairman in his closing comments. A better discipline on the part of some of the speakers (and perhaps a harder line by the Session Chairmen) was clearly needed.

10.4 Presentations were, on the whole, good. In some cases, the visual aids used (e.g. typewritten viewgraphs) were not very visible and some speakers without good English were difficult to follow. There was perhaps also a tendency to try to cram too much into the allotted speaking time – earlier availability of papers could help in this and perhaps enable speakers to select the major points of their papers for presentation. In a few cases, for example, the speaker merely read his printed paper. The translators seemed on occasion to have had little time beforehand to become familiar with the papers – interpretation thus suffered – and authors should provide material to the interpreters well ahead of their presentations.

10.5 As to the local organisation of the meeting, everything went very smoothly and efficiently. The AGARD staff and those of the Air War College involved are to be congratulated. One of the main advantages of meetings of this sort is the opportunity they give of meeting other experts in the field and often the best results are obtained outside the meeting room. During the day there were adequate opportunities to bring this about; the scattering of delegates over many hotels, inevitable in a city like Florence, still in the tourist season, perhaps restricted evening discussions and social meetings.

10.6 It is difficult to assess audience reaction to the symposium. Only three questionnaires were returned and even if these were generally favourable the sample is hardly a fair one. The comments made are, however, incorporated here and in the next section of this report.

10.7 As has been noted in previous sections, two of the papers were classified NATO Confidential for no very apparent reason. This had the effect of unnecessarily complicating the local arrangements and care must be taken at future meetings not to over-classify material.

10.8 It was surprising to find no ESA representation at the symposium. Perhaps the military label was to blame but it appeared that NASA had no qualms about attending. It was perhaps also surprising that the USAF Space Division was not present.

11. CONCLUDING COMMENTS ON TECHNICAL CONTENT

11.1 Technical Balance

11.1.1 The Programme Committee put together a reasonably good set of papers – bearing in mind their dependence to a large extent on what is offered to them. The balance was good, covering most of the areas of current concern in control system design. There was a nice mixture of down to earth papers (if such a phrase can be used in relation to a space symposium) and of high-flown theory which will need time to digest. Perhaps more would have been desirable on the characteristics of the sensors and actuators needed in future programmes – the session on control components did little more than discuss such as were currently available and well known.

11.1.1 In a few cases, the authors contented themselves with reporting the results of work done and a look into the future would have been desirable, with an identification of new issues that had emerged and of further research requirements. In other cases, work was clearly at an early stage but it was useful to hear progress reports. These comments do not, of course, uniformly apply and there were adequate looks at the future – NASA's views on needs were valuable for example. The highly theoretical papers of Session V certainly raised some controversy and it will be interesting to observe how the gap between the theoreticians and the practical control engineers varies with time. The comment made during the Round Table Discussion that, at the first IFAC meeting in 1956, an objective that emerged was to close the gap between theory and practice seems again relevant.

11.2 Technical Issues

11.1.1 It was apparent that a major area of immediate concern in control system design was how to deal with the interaction between the spacecraft structure and the control system itself and the need to establish a sound basis of theory was underlined. The large spacecraft of the future with their many flexible appendages, space stations assembled

in orbit element by element with consequent changes in dynamic behaviour and the increasing accuracy demanded of attitude control all present problems of great complexity to the control system designer. System modelling and simulation will be necessary but it was also to be noted that much emphasis was placed on the need for flight tests of real systems. The various issues were well brought out.

11.2.2 It was interesting to note the current European concern with the problems of rendezvous and docking in orbit, an area where the USA has already reached a high level of competence. European work on this topic is clearly at a very early stage. A start has been made, however, and it is clear that this will remain a strong research area for some time.

11.2.3 The importance of software also emerged. The microcomputer presents many opportunities to the control system designer but it brings its own problems too. There is the reliability of the computer itself but perhaps more important is the 'reliability' of the software - the need, for example, to be sure that the programs will perform all the necessary functions in the prescribed environment and that all eventualities, including failure situations, have been catered for. There is the need, too, to be able to modify on-board software from the ground as the dynamic environment changes. All this underlines the need for adequate spacecraft models: quality control is every bit as important with software as with hardware. In-flight testing of software is also available and may indeed be desirable, just as in the case of some hardware - there was an interesting presentation on this topic from which there emerged that a spacecraft mathematical model is not always a proper substitute for the real thing.

12. MILITARY IMPLICATIONS

12.1 As was noted earlier, there was little or nothing of a specifically military nature discussed at the symposium perhaps, as was said, not surprising since as far as guidance and control is concerned, civil and military systems have requirements in common and also in view of the fact that European NATO members have had little involvement with space systems for military purposes. Some greater input from the USA in this area might have been expected. The papers presented, of course, have relevance to military systems but it would have been desirable to have heard something from the military customers in each NATO country, of their views on military space systems and of the requirements they foresaw. This could be a useful element in any future meeting held by AGARD on space.

13. RECOMMENDATIONS

13.1 In any future AGARD meetings on space-related subjects it would be desirable, as has just been said, to hear the views of the military planners, within NATO organisations and those of the participating nations, on future requirements for space systems for military purposes. It might well be said that none of the papers given in Florence would have been out of place at any unclassified symposium and an injection of military thinking could help in future events of this sort.

13.2 It would be wrong to make any specific proposals as to areas where research is necessary and indeed difficult to visualise how AGARD could bring influence to bear. As evidenced by the symposium, the major problem areas in control system design are generally well recognised and are being investigated. In Europe, of course, most of the effort is being supported by civil funds and conceivably AGARD could influence the military planners of the member states (the US apart) to think harder about their needs for space systems. Or is this an area to be left entirely to the USA?

13.3 In paragraph 10 a number of comments were made on the symposium arrangements and perhaps it suffices to underline a few of them:

- (a) it would be better if papers were available for study before the meeting; this could help in streamlining presentations and in allowing longer time for discussion;
- (b) there is a need for greater discipline on the part of speakers to stick to their allotted time and to use visual aids that are visible;
- (c) over-classification of papers should be avoided or alternatively, if the meeting is to be classified, then classified material should be presented.

CONTROLS TECHNOLOGY THE KEY TO FUTURE SPACE SYSTEMS

by

Robert J. Herzberg
Lockheed Missiles and Space Company
062-07 B 104
1111 Lockheed & Space Co.
Sunnyvale CA 94086
USA

INTRODUCTION

When I was asked to stand in for Del Williams as your keynote speaker, I was honored to be asked to participate in what I consider a timely and significant symposium directed to issues of major concern for future space programs. I am convinced that the progress in space-technology which we have experienced in the past years is due in no small part to gatherings such as this. The evidence of this progress is apparent in the current generation of satellites and spacecraft, and the technical community which you all represent can justifiably take pride in the accomplishments to date.

As impressive as these accomplishments are, however, future space initiatives pose even more significant challenges which will require well planned, innovative technology development programs. In a number of areas major advances are needed if we are to attain our goals in space, and progress in these areas is pivotal to ensuring engineering readiness for the next generation of spacecraft. I'm sure it's no surprise to this audience that one of these critical areas is spacecraft guidance and control.

Since my primary responsibility to the Space Systems Division of the Lockheed Missiles and Space Company is the planning and overseeing of our technology development programs, I am particularly interested in the advances in controls technology that will be presented and discussed in your symposium proceedings. In turn, I would like to share with you some thoughts and observations regarding the role that your technology can play in enabling future space missions that are presently in the preliminary and advanced planning stages. These observations are derived both from my involvement with the technology planning activities at Lockheed and from my experiences as a member of NASA and U.S. Air Force technical committees.

At the outset, I would like to point out that I am not an expert in guidance and controls technology, and my remarks will not address the exciting new technology developments that will be ably presented by the speakers that will follow me. Instead, my comments will be from the point of view of one who is concerned with the broad range of issues relating to technology readiness for future space programs, many of which have vital military significance for the free world. From this perspective, I will discuss the role advanced guidance and control technology can be expected to play in enabling future space missions and the perceived difficulties involved in meeting the challenges. It is my hope that these remarks will help underscore the significance and timeliness of this symposium and the importance of your work in preparing the way for the next generation of space vehicles.

OVERVIEW

Although the space age is barely two and a half decades old, we have seen the growth from tiny, rigid bodies capable of transmitting meaningless electronic signals to complex, multi-bodied vehicles capable of performing a vast array of important science, military, and commercial applications missions. The evolution of guidance and control technology to support these missions has been correspondingly dramatic, and this evolution has been the subject of a number of technical papers in recent years. As impressive as these accomplishments have been, it is important to recognize that the progress has been primarily in the area of applications rather than theory. In fact, several noted authors have taken the position that spacecraft guidance and control is so well established that it no longer constitutes a field for research, with the strong implication that future space programs can consider this area as a "flight-ready" technology rather than one requiring major attention for future development.

Is this, in fact, the case? Can those who have the responsibility for planning future space missions consider guidance and control issues to be routinely handled by normal technological developments, or will they, instead, be seriously constrained in their options for future spacecraft configurations if they rely on extensions of our present practices? These questions, which are fundamental to the theme of this symposium, are of the greatest significance to the space mission planners, and they have been receiving major attention at the highest levels of U.S. space program management. NASA recently formed a special committee to report on these issues, and the subject continues to have high-priority NASA attention. I believe you will hear something of this in the following presentations. The U.S. Air Force has an on-going effort to define needed technology advancements, and the Military Space Systems Technology Model (MSSTM), which defines these requirements, has a substantial section devoted to guidance and control issues. Also, I believe I'm on safe ground in stating that all major U.S. aerospace contractors are watching this area with great interest as part of their in-house technology development program planning. My following remarks will be directed to summarizing the general findings of these various planning groups relative to the needs for advanced controls technology, and the perceived status of the technology to meet these needs.

FUTURE NEEDS

Perhaps the most visible example of future spacecraft trends affecting controls technology is the Space Telescope (ST). The size of the satellite and the precise accuracy requirements posed substantial controls challenges, and its operation will be a milestone in fine-pointing control of spacecraft. The AXAF (Advanced X-ray Astrophysical Facility) program, which is in an advanced planning stage, will benefit from ST experience, but will provide its own challenges for controls technology. Both of these vehicles typify the trend to large articulated spacecraft possessing flexible appendages. Also, for future planetary missions using solar-electric propulsion, the major portion of the spacecraft structure will be solar arrays, and effective handling of the distributed flexibility of these systems will be a significant controls issue.

With the advent of the Space Transportation System (STS), it is now possible to consider very large spacecraft assembled in space using multiple STS launches. Although a practical solar-power space station is not realizable in the near future, considerable solar array area is required for a number of future space missions which are in present planning stages. Such missions can be expected to have flexible-body modes close to or within the necessary controls bandwidth, and they will require new controls approaches. Finally, the NASA Space Station initiative is in its formative stages, but is a strong motivation for controls technology development. Although the final configuration is still evolving, it seems clear that advanced controls techniques will have to be considered, and that such techniques may significantly enhance the design.

The two most significant controls technology issues that have emerged from the technology needs assessments of these missions are related to the potential need for controls strategies that actively suppress flexible-body structural responses and for autonomous controls operation. The latter issue is of extreme importance to a number of future space programs, especially in the military sphere, and I will comment further on this a little later. However, the issue causing greatest concern as regards new technology development is whether or not present controls approaches will be able to satisfy the extremely demanding accuracy and agility requirements of future space missions, some of which also require very large spacecraft that will be considerably more flexible with respect to the controls bandwidth than has been the case to date.

For purposes of brevity, I will refer to controls approaches that include the active suppression of elastic mode responses as CSI (Controls-Structures-Interaction) technology in this discussion. While this technology has considerable development history, this has been almost entirely in a research/university context. No present spacecraft system uses this approach, and relatively little work has been done to transition CSI technology to engineering applications. In the U.S. a significant start has been made via a government sponsored technology initiative termed ACROSS (Active Control of Space Structures), which is funded by the Defense Advanced Research Programs Agency (DARPA), but much work remains to be done. As a technology planning issue, then, it is clearly of great importance to assess the need for this advanced controls approach for future spacecraft and to realistically identify the development timelines for remedying any deficiencies.

One of the more intensive attempts to evaluate the need for CSI technology for future U.S. space programs was the NASA committee activity I earlier mentioned. This considered both NASA and military candidate missions. Although determining the level of need for CSI approaches requires a reasonably complete preliminary design process, which was not available for most of these future missions, the committee was able to conclude that a number of major future NASA and military space initiatives would face serious design constraints if CSI is not available as a controls option. The most stringent controls requirements were related to resolution and pointing accuracies of primary payload sensors, some of which require very large apertures.

A useful technique for assessing when the combination of size and accuracy requirements for these spacecraft will result in a probable need for advanced controls technology is to consider the parameter D/f , where "D" is the aperture diameter and "f" is the lowest modal frequency of significance to the pointing and configuration control of the system. If this is plotted as a function of operating wave-length for various missions, those having a D/f greater than approximately 100,000 can be considered in the CSI technology category. This approach, developed by Dr. K. Soosaar of Draper Laboratories, is illustrated in Figure 1, which identifies 33 missions presently in serious planning stages. While this should be viewed as only a rough assessment of need in any specific case, the overall picture clearly supports the conclusion that the CSI option is a major concern for many future space missions.

The second major area of concern regarding guidance and control technology for future space programs that has emerged from these planning studies is the need for greatly enhanced capabilities for autonomous operation. The ability of spacecraft to maintain on-board knowledge of its orbit parameters and to perform its navigation and station keeping functions for long periods without contact from ground stations is key to successful performance of a number of important missions. This requirement has been largely related to military needs for survivable space defense assets, but there are also major cost savings associated with reducing the very sizable expenses incurred by ground operations related to constant monitoring of spacecraft functions and determination of orbital parameters. The Jet Propulsion Laboratory of NASA has taken the lead in defining autonomy capability levels for spacecraft operations and is presently working with the

U.S. Air Force to help focus the needed technology developments in this area. They have defined a graduated scale of 10 autonomy levels for assessing spacecraft needs. Using this scale, an overall assessment of future requirements for our defense satellites shows a major need for improvements in the controls area, especially in on-board orbital parameter determination, fault tolerance, and orbit correction decision making.

TECHNOLOGY ASSESSMENT

One of the most difficult problems faced by those responsible for our space programs is the objective and accurate assessment of the engineering technologies necessary to field complex future space systems. Errors in judging the ability of our technical community to meet these needs can obviously have the most serious consequences. Unfortunately, a number of past projects have had to be cancelled, or were completed with large over-runs and reduced performance, due to inadequate technology assessments prior to their initiation. The major issue is not technical feasibility, although this, of course, must be established, but rather whether or not the technology will be "flight-ready" when it is needed. The term "flight-ready" in this context implies (1) minimal risk compared to other approaches, (2) cost effectiveness, (3) a reasonable level of proven capability by aerospace contractors, and (4) established procedures for developing associated performance and testing specifications for spacecraft systems.

As an example of this assessment process, the following questions must be carefully considered:

- Is the theory well developed and are the necessary computational tools available?
- Has the technology been adequately tested on previous spacecraft, or by laboratory experiments?
- Has the technology been reduced to engineering practice? (i.e. specifications and criteria)

While obtaining such an assessment may appear to be straight-forward to those not involved in space technology management, in practice quite the opposite is true. The program office people who are tasked with planning future projects may tend to underestimate technical risks, either because of a lack of technical expertise or due to their understandable enthusiasm to "sell" their projects. On the other hand, the technologists frequently tend to be over-optimistic regarding technical risks and often do not understand the engineering difficulties associated with converting their technical contributions into flight hardware that can be delivered on-time and within cost.

With these comments in mind, I would now like to briefly discuss the assessment of controls-structure interaction technology that was earlier identified as a major concern for future space missions. A first observation regarding CSI technology is that it cannot reasonably be viewed as a simple extension of existing practice. It's application to future spacecraft will require substantial changes in the way the controls subsystem is handled, and a number of technical issues must be resolved before it can be considered flight-ready. Some deficiencies are evident in each of the three assessment categories I earlier mentioned. The methodology area is closest to readiness, although even here some questions need resolution. In the experimental base and practice and criteria areas much work needs to be done. Five of the most important sub-categories of CSI technology that need work are:

- Analytic modelling accuracy and model reduction criteria
- System identification procedures
- Robustness criteria
- Development of sensors and actuators
- Digital implementation

Each of these areas is of particular concern, and technology development activities are needed if we are to realize the potential of this new controls approach for future spacecraft.

Perhaps one of the most important aspects of this new approach is the need to recognize that its success will require the cooperative efforts of several disciplines in a much more effective manner than has been the case in the past. The practical convergence of controls, structures, and computer technologies will be necessary if CSI applications are to become a reality (Figure 2). Multi-disciplinary research and development programs, improved university curricula, and new approaches to industrial technical management are all important. This symposium and the efforts of this technical community will be key elements in this evolution, which will play a major role in preparing the way for the next generation of space systems.

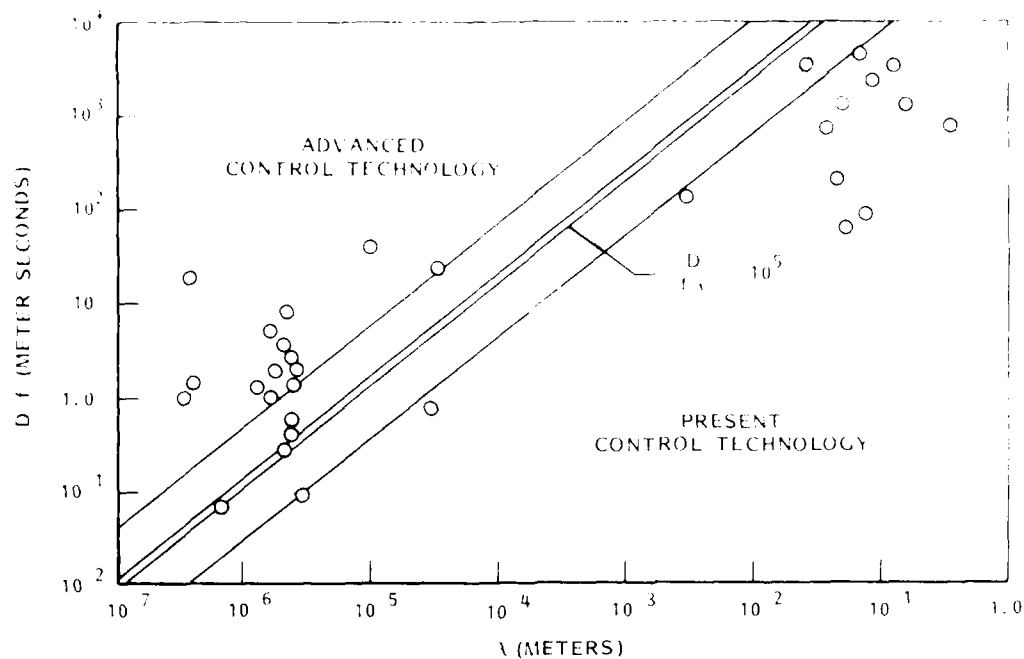


Figure 1 CONTROL TECHNOLOGY REQUIREMENTS FOR FUTURE U.S. SPACE MISSIONS

REQUIRES A NEW APPROACH TO TECHNOLOGY INTERFACES

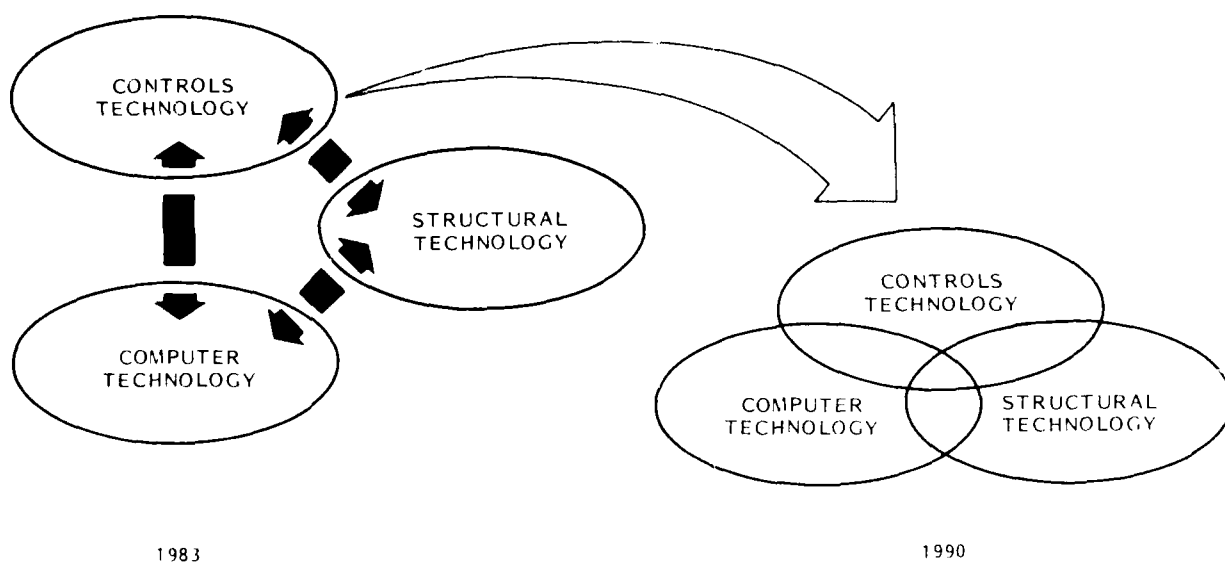


Figure 2 CONTROL OF FLEXIBLE SPACECRAFT

SPACECRAFT CONTROL RESEARCH AT NASA

John B. Dahlgren
Program Manager, Space Controls and Guidance R&T
Office of Aeronautics and Space Technology, National Aeronautics
and Space Administration, Washington, D.C. 20546, USA

Lawrence W. Taylor, Jr.
Manager, Spacecraft Control Branch
NASA Langley Research Center, Hampton, VA. 23665, USA

SUMMARY

Future missions in space will require controlling spacecraft which are both large and flexible. The limited inherent damping and the uncertain and changing dynamic characteristics of many of these vehicles, such as manned space stations and large antennas, will revolutionize spacecraft control requirements. In preparation for the time that such control systems are required, considerable research and technology development is necessary. A program is in place at NASA for the development of active control technology to support major initiatives for space station and advanced spacecraft. A number of key control technology program needs are cited in the paper as required for these and other future NASA missions together with an integrated controls/structures technology flight experiment to demonstrate and validate technology for large flexible structures.

INTRODUCTION

The objectives of NASA's space technology long range plan are shown in Figure 1. As indicated in the figure the importance of space R&T must be established and strengthened to ensure the timely provision of new concepts and advanced technologies for the U.S. civil and military space activities. New initiatives are being planned to provide the capability for in-space testing through space flight facilities and experiments, and to develop the technology to support space station and advanced spaceframes. Also a new program has begun to increase university research and small business innovative research which address important NASA scientific or engineering problems and opportunities. Before moving into details of spacecraft control research at NASA which is associated with the theme of this technical symposium on "Guidance and Control Techniques for Advanced Space Vehicles" we will first look at the specific objective for the NASA space controls and guidance R&T program shown in Figure 2.

OBJECTIVES

- ESTABLISH IMPORTANCE OF SPACE R&T TO ATTAINMENT OF NATIONAL SPACE POLICY OBJECTIVES
- PROVIDE U.S. R&T CAPABILITY BY MAINTAINING NASA CENTERS IN POSITIONS OF UNDISPUTED EXCELLENCE IN CRITICAL SPACE TECHNOLOGIES

INITIATIVES: IN-SPACE FLIGHT FACILITIES AND EXPERIMENTS

- STRENGTHEN NASA'S SPACE R&T PROGRAM TO INSURE THE TIMELY PROVISION OF NEW CONCEPTS AND ADVANCED TECHNOLOGIES FOR THE U.S. CIVIL AND MILITARY SPACE ACTIVITIES

INITIATIVES: SPACE STATION TECHNOLOGY
SPACEFRAME TECHNOLOGY

- ASSURE BALANCED PARTICIPATION IN SPACE R&T PROGRAM BY NASA CENTERS, GOVERNMENT AGENCIES, UNIVERSITIES, AND INDUSTRIAL RESEARCH ORGANIZATIONS

INITIATIVES: UNIVERSITY RESEARCH

Figure 1. Space technology long range plan objectives

SPACE CONTROLS AND GUIDANCE R&TSPECIFIC OBJECTIVE

- o TO DEVELOP
 - o DESIGN CONCEPTS, ANALYSIS, AND TESTING TECHNIQUES
 - o ADVANCED COMPONENTS
- o FOR CONTROLLING AND GUIDING EARTH-ORBITING SPACECRAFT INCLUDING
 - o LARGE ANTENNA
 - o PLATFORMS
 - o PAYLOADS AND EXPERIMENTS
 - o SPACE TRANSPORTATION SYSTEMS

Figure 2. Space controls and guidance specific objective

This objective is quite general but covers a number of thrust objectives shown in Figure 3. Although the thrust objectives are focused toward a particular area or mission a substantial portion of the resulting technology will have general applicability to all areas.

CONTROL AND GUIDANCE R&T THRUST OBJECTIVES

- o GENERIC
TO DERIVE NEW CONCEPTS AND APPROACHES TO EFFECTIVE CONTROL OF LARGE DYNAMIC STRUCTURES.
- o TRANSPORTATION
TO DEVELOP ADVANCED CONTROL AND GUIDANCE CONCEPTS FOR CURRENT AND FUTURE TRANSPORTATION SYSTEMS WHICH ARE MORE TECHNICALLY EFFICIENT AND WHICH LEAD TO MORE COST AFFORDABLE FUTURE SPACE TRANSPORTATION MISSIONS.
- o SPACECRAFT
TO DEVELOP ADVANCED CONCEPTS AND TECHNIQUES FOR CONTROLS, PRECISION POINTING, AND STABILITY AND GUIDANCE AND NAVIGATION OF LARGE ANTENNAS, PRECISION REFLECTORS AND ADVANCED EARTH ORBITAL SPACECRAFT.
- o SPACE STATION
TO ADVANCE THE UNDERSTANDING AND STATE OF THE ART IN POINTING AND STABILIZATION OF LARGE SYSTEMS CONTROLS, MULTIBODY GUIDANCE AND CONTROLS, AND ADVANCED GUIDANCE AND CONTROL DEVICES.

In the following sections, highlights will be given in each of the thrust areas. In addition the issues of control structure technology and the need for space flight experiments to demonstrate and validate this interactive technology will be discussed.

GENERIC RESEARCH

In the research area NASA is currently investigating a number of theoretical areas which have been identified as being deficient for applications to large, highly flexible spacecraft and requiring further development and validation through experimental hardware testing. Current theoretical deficiencies identified from technology surveys and systems studies are shown in Table 1. Most of the indicated theoretical areas are being investigated through a combination of university grants and NASA center research programs. Another area of research begun in 1983 is in the applications of distributed system theory to control of large space structures. The term distributed systems as used here has two meanings. One is control theory for distributed parameter system modeling and the second is the theory for systems requiring spatially distributed multi-point sensing and actuation. Many large space structure configurations being considered in the NASA mission model will require multi sensors and actuators distributed in some fashion throughout the structure. Therefore, one must consider the spatial distribution of sensors and actuators whether treating the problem from the finite elements or continuum model approach.

Table 1. Current Theoretical Deficiencies in Control of Large Flexible Spacecraft

Fundamentals of Control of Highly Flexible Systems	Adaptive/Learning Control	Optimal Management of Distributed Control Systems
Modern vs Classical	Identification of Most Important Parameters for Control	Fault Modeling
Linear Quadratic Gaussian	Sensitivity of Real Time Models	Failure Detection
Decoupled Control	Parameter Identification	Self-Reorganization
Computer Aided Design	Real Time Automated Control Laws	Analytic Redundancy
Dual Level Control	Parameterization of Shape Configurations	Optimal Reduced Capability Performance Systems
Modeling for Analysis	Criteria for Closed Loop Adaptation	
Parameter Sensitivity		
Robust Control Laws		

In adaptive control research, techniques are being investigated for their future application to control of large space structures where uncertain or changing parameters of the system may destabilize conventional control system design. In one ground experiment using a large flexible hanging beam and realistic hardware the application of a state space adaptive filter/controller was demonstrated which accounted for an intentional error in its 3rd eigenvalue or modal frequency. Using colocated control sensors and actuators, and a high fidelity structural model and feedback controller with a Kalman filter and optimal control laws, beam vibrations were rapidly damped following an impulse disturbance at the free end. With the intentional error of approximately 20% in the 3rd mode replaced in the model, the control system was shown to be unstable for the same initial conditions. However after parameter identification and adaptive estimation techniques were added to the feedback loop, the system stabilized after one and one-half cycles of vibration. Real time computational requirements were handled by a Motorola 6502 microprocessor operating at 1 MHz.

In a related area of research adaptive identification for the dynamics of large space structures is being investigated using lattice filters. The approach illustrated in Figure 4 is for the problem of identifying the structural dynamic characteristics of basic structures. The lattice filter provides an algorithm based on least squares estimates

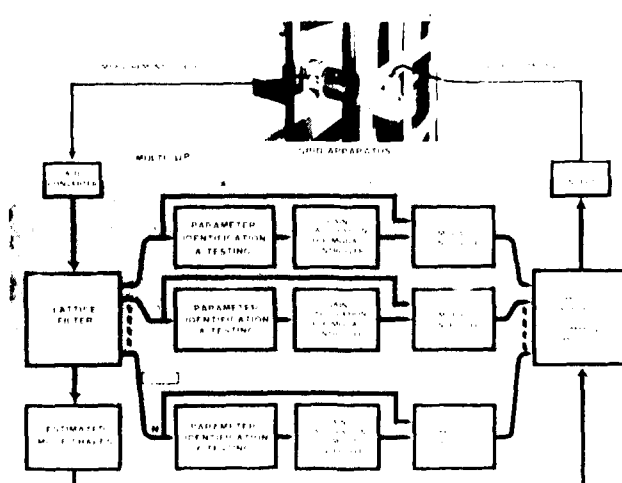


Figure 4. Distributed adaptive control

for generating an orthonormal basis for the measurement data using measurement samples. Using this approach a modal system description is avoided and hence the a priori specification of model order (i.e., number of modes requiring control is not specified). Instead the systems order is recursively determined on-line, as would be required in flight, along with its parameters. The use of recursive lattice filters are widely used in adaptive signal processing. The next step in this research will be the application of lattice filters to two dimensional structures.

Large space systems (LSS) control technology

For control of future large space systems a number of technology needs have emerged. Large as used here is a relative term for those systems which are larger than systems flown to date and which carry high performance requirements. Many of these needs serve as the focus for principal activities, shown in Figure 5, in the space station and spacecraft programs. If we look at the NASA LSS mission overview, shown in Figure 6, we have a number of mission classes for which the objectives are quite distinct from past missions. Space station will be characterized with designs providing for evolutionary growth, low life cycle cost, and user accommodation. The NASA large space antenna missions can be categorized into three major classes: radio frequency (RF) antennas, large segmented reflectors, and LSS flight experiments. The large RF antenna class includes near-term communications missions, such as the Land Mobile Satellite System, radiometry missions, and the advanced communications missions. The segmented reflector systems will require multifaceted solid-panel reflectors deployed on supporting structures to meet the requirements for IR and submillimeter astronomy. Flight experiments to validate LSS technologies may involve fundamental structures with associated control and build up to a fully deployable large space operational antenna. In the following sections control technology needs and requirements for these missions will be highlighted.

LARGE SPACE SYSTEMS CONTROL TECHNOLOGY

- ANTENNA CONTROL
- PLATFORM/STATION CONTROL
- CONTROL OF DISTRIBUTED PARAMETER SYSTEMS
- SYSTEMS IDENTIFICATION
- FIGURE CONTROL
- EXPERIMENTAL VALIDATION
- MODEL ERROR ESTIMATION
- ADAPTIVE CONTROL
- ANTENNA CONTROLLER DESIGN BASED ON RF
- ADVANCED GUIDANCE AND CONTROL COMPONENTS
- FLIGHT EXPERIMENT DEFINITION AND DESIGN
- INTERVEHICLE CONTROL

Figure 5. Principal activities in program

CLASS	OBJECTIVES
<ul style="list-style-type: none"> • PLATFORMS AND SPACE STATIONS <ul style="list-style-type: none"> • UNMANNED SPACE PLATFORMS • MANNED SPACE STATION 	<ul style="list-style-type: none"> • MULTIPLE USER SPACE/EARTH OBSERVATIONS • TRANSPORTATION NODE TO HIGHER ENERGY LEVELS • ASSEMBLY, CONSTRUCTION • 0-g PROCESSING, MANUFACTURING
<ul style="list-style-type: none"> • LARGE ANTENNAS 30-100 m <ul style="list-style-type: none"> • MESH DEPLOYABLES • TRUSS • ELECTROSTATIC MEMBRANE, etc. 	<ul style="list-style-type: none"> • COMMUNICATIONS • RADIOMETRY/EARTH RESOURCES • RADIOASTRONOMY
<ul style="list-style-type: none"> • SEGMENTED REFLECTORS 10-30 m <ul style="list-style-type: none"> • LARGE DEPLOYABLE REFLECTOR 	<ul style="list-style-type: none"> • IR, SUBMILLIMETER ASTRONOMY
<ul style="list-style-type: none"> • LSS FLIGHT EXPERIMENTS 	<ul style="list-style-type: none"> • VALIDATION OF ENABLING TECHNOLOGIES • ESTABLISH REQUIRED FLIGHT DATA BASE

Figure 6. NASA LSS mission overview

Space station

A number of space station concepts have been evolved in recent times as shown in Figure 7. The figure on the right shows space station in its most current conceptual approach which is characterized by evolutionary and modular growth. There are many physical factors which must be considered in such an evolutionary process for space station. Some of these important factors are shown in Figure 8. All of these physical factors will seriously impact the station's control systems performance and operation in its initial configuration and potentially much greater as the station evolves into more advanced configurations with additional modules and provisions for vehicle servicing

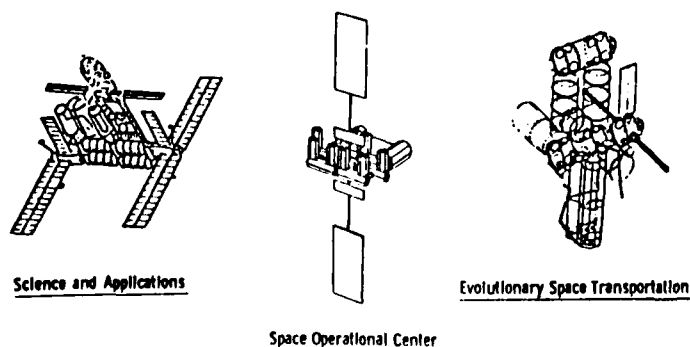


Figure 7. Developing space station concepts

1. INCREASING ORBITAL MASS
2. MIGRATING CENTER-OF-GRAVITY LOCATIONS
 - o STATION BUILD-UP
 - o CHANGE-OUT OF MAJOR ELEMENTS (SOLAR ARRAYS, RADIATORS)
3. VARYING AERODYNAMIC AND GRAVITY GRADIENT TORQUES
 - o ORBIT RE-BOOST
4. CHANGING STRUCTURAL RIGID BODY AND FLEXIBLE MODES
 - o STATION BUILD-UP
 - o DOCKED ORBITER/ORBIT TRANSFER VEHICLE (OTV)
 - MASS TRANSFER OF FUEL/SUPPLIES
5. ASTRONAUT/CONSTRUCTION MOVEMENTS

Figure 8. Physical factors of the evolutionary process

(e.g., orbit transfer vehicle). Of those factors listed, the factor of changing structural rigid body and flexible modes imposed during the station build-up and docking operation, with the station places new and unique requirements on space control systems. In order to assess the evolutionary changes required in a control technology development program which addresses large space systems a controls technology roadmap was formulated as shown in Figure 9. The chart illustrates technology readiness for LSS as projected by the long range plan for control technology development. Currently we see the state-of-the-art for control of large structure as represented by technologies applied to Skylab, Space Shuttle, and recent spacecraft missions. This control has been based on structures being modeled

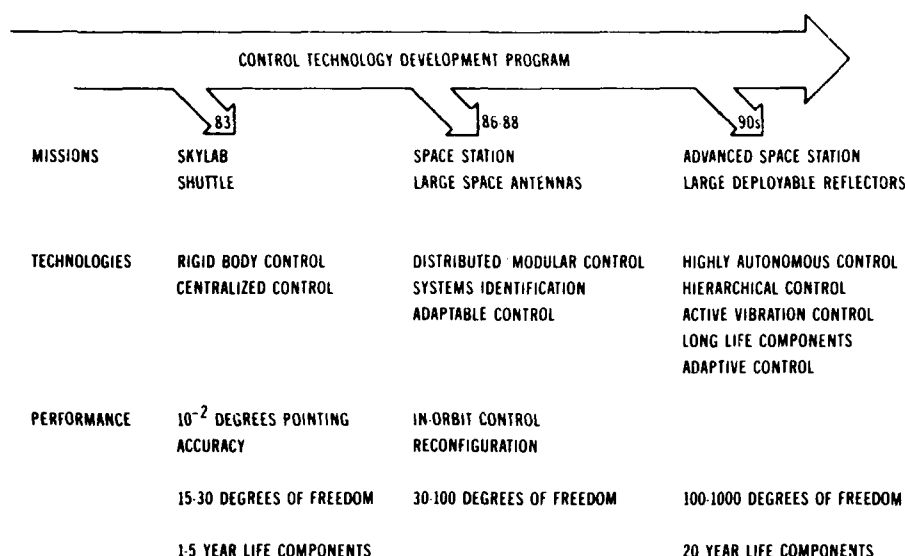


Figure 9. Control technology evolution

as a rigid body with flexible appendages. Centralized controllers with appropriate notch filters to minimize coupling with flexible modes in the control bandwidth was adequate to meet the requirements. For the initial space station such a conventional approach for control may also be adequate depending on the dynamics and user accommodation. After a few or more modules are joined in space, the structural dynamics will interact with the control dynamics in such a way that modern control theory with distributed control will be highly desirable if not mandatory. The technology readiness date for the initial station is in the 1986 to 1988 time period. A control system block diagram for the initial station might take the form of that shown in Figure 10 where distributed modular control would be combined with some level of systems identification and adaptive capability. This early form of adaptive control we refer to as adaptable control to accommodate changing configurations and operational requirements. Here we might use gain scheduling techniques for control in the absence of systems ID, and tracking notch filters combined with systems ID for fine tuning the adaptive controller. As we move from adaptable forms of control to fully adaptive, which may use an expert knowledge base for adaptive problem solving, we might envision the hierarchy as shown in Figure 11. As we move from the top to the bottom we increase the autonomy of the system but with some increase in complexity. It is felt that some moderate level of these technologies would provide the baseline in control

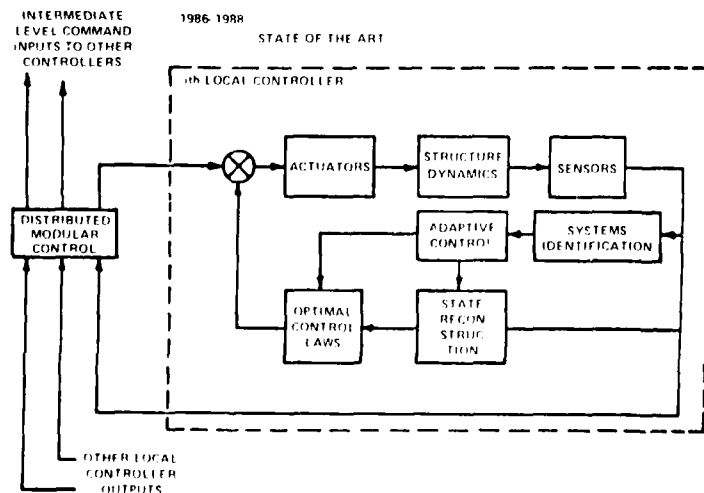


Figure 10. Advanced control technology integration

- Gain Scheduling Controller
 - Configuration and operations adaptive.
- Systems Identification
 - Mass Properties Tracking
 - Disturbance Measurement
 - Flexible Structure Behavior
- Adaptive Compensation Filters
 - Tracking Notch Filters
 - Disturbance Adaptive Estimators
- Model Reference Adaptive Control
 - Adaptive State Observer
 - Adaptive Parameter Identifier
 - Adjustable State Regulator
- Self Tuning Regulators
 - Adaptive Predictors
 - Stochastic Self Tuning Controllers
- Distributed Adaptive Systems
 - Multi Input, Multi Output
- Adaptive Problem Solving System
 - Expert Knowledge Base Plus Distributed Adaptive Control

Figure 11. Mission adaptable to adaptive control hierarchy

for an evolutionary, modular growth station. For the 1990's the objective of the controls program is to provide technology options and opportunities such as those listed in Figure 9 for advanced space station application.

SPACECRAFT TECHNOLOGY

Presently the spacecraft controls technology program deals with the requirements for large antennas and submillimeter wavelength, and lower, astronomy spacecraft. Control technology requirements for these systems are shown in Figures 12 and 13. Control technology requirements in Figure 12 fall into the five groups shown on the left with technology level of advancement dictated by the application areas shown across the top. Figure 13 illustrates the broad spectrum of pointing and figure control requirements spanned by large antennas. From the near-term communications antennas and radiometers to the VLBI and advanced communications antennas (such as the RF Orbiting Deep Space Relay Station), the surface and pointing requirements are at least 1 order of magnitude apart. From the advanced RF systems to the IR and optical systems (such as the laser orbiting deep space relay station), the requirements are another 3 orders of magnitude tighter representing a giant step in technology advancement.

	NEAR TERM COMMUNICATIONS	ADVANCED COMMUNICATIONS AND RADIOASTRONOMY	IR, SUBMILLIMETER ASTRONOMY
DISTRIBUTED SENSING AND ACTUATION	15-30 DMF	10-60 DMF ACTIVE DAMPING	~500 DMF ACTIVE DAMPING
SYSTEM IDENTIFICATION	INITIAL SYSTEM DYNAMICS, PERIODIC CALIBRATION	INITIAL SYSTEM DYNAMICS, PERIODIC CALIBRATION	REAL TIME SYSTEM DYNAMICS AND DISTURBANCES
ADAPTIVE CONTROL	GROUND-BASED REPROGRAMMING	GROUND-BASED REPROGRAMMING	REAL TIME GAIN ADJUSTMENTS
FIGURE DETERMINATION AND CONTROL	INITIAL ADJUSTMENT, PERIODIC CALIBRATION TO 1/2 mm	INITIAL ADJUSTMENT, PERIODIC CALIBRATION TO 1/20 mm	ACTIVE TO 1/2 μ m AT 1 Hz
LONG-LIFE ACTUATORS AND SENSORS	FIGURE SENSORS TO 1/2 mm AT 5 Hz	PRIME-MASS DAMPERS, FIGURE SENSORS TO 1/20 mm AT 5 Hz	HIGH RESOLUTION, LOW NOISE WHEELS, PROOF-MASS DAMPERS, FIGURE ACTUATORS, FIGURE SENSORS TO 0.1 μ m, LOS TRANSFER SYSTEM TO 0.05 μ m

Figure 12. Antenna control technology requirements

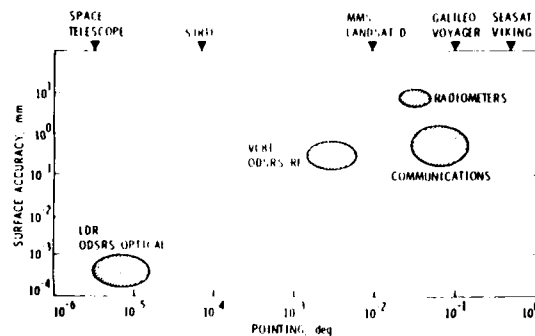


Figure 13. Antenna surface accuracy vs pointing requirements

Land mobile satellite service

The Land Mobile Satellite Service (LMSS) is a representative example within the large RF antenna mission class. The LMSS is one of the early large antenna applications which has been studied in considerable detail by a joint team of JPL/NASA and industry technologists. The LMSS is a multibeam communications mission utilizing a geosynchronous-orbit-based large antenna for providing telephone services to mobile users in the continental United States. Two candidate LMSS configurations have been studied. One is a wrap-rib design consisting of a 55-m-diameter mesh reflector and a large 8 x 11-m RF feed. These are connected by two booms 34 and 80 m in length. The second configuration is a hoop-column design consisting of a 118-m-diameter mesh reflector, an 88-m mast, and four feed arrays of 4 x 4 m each. Each of the systems weighs about 10,000 lb and have moments of inertia of 10^6 to 10^7 kg-m².

The principal control system objective for a RF antenna such as LMSS operating at a frequency of 0.87 GHz is to point the RF beam(s) to the desired target(s) within prescribed pointing and jitter errors (typically 0.020° to 0.002°) while maintaining overall system alignment and figure to insure the desired RF performance (main beam gain, low sidelobes, etc.). Achieving the desired antenna pointing performance on a flexible structure of 50 to 100 m while maintaining the static and dynamic figure within an envelope of 0.5 to 5 mm represents a substantial challenge to the technology. The two LMSS antenna configurations under current study are shown in Figure 14 under dynamic disturbances. The distortions and displacements have been amplified to illustrate the pointing problem. The major contributors to RF pointing error are spacecraft attitude errors, feed displacements and rotation, and dish attitude and deformations.

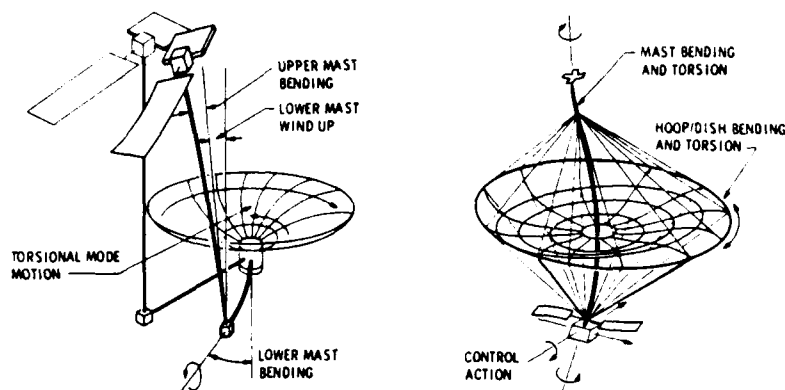


Figure 14. Antenna pointing problem

Hoop-column antenna control design analysis

A control-design analysis is under development for a hoop-column antenna system. A NASTRAN model of a 122-m-diameter hoop-column antenna is being used as the basis for the design. The schematic sketch, Figure 15, shows possible locations for the two types of controllers employed in the simulated studies. Control moment gyros and/or reaction

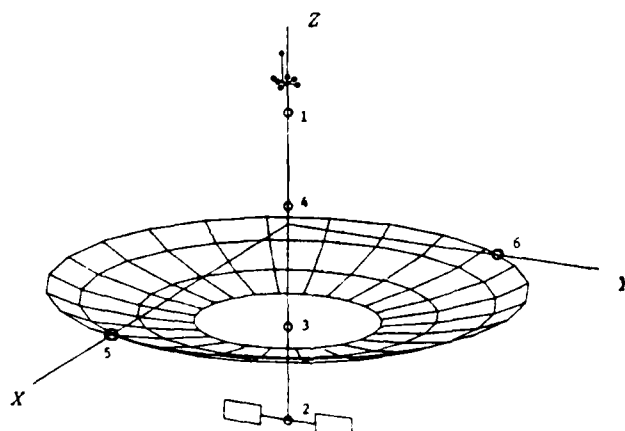


Figure 15. Hoop-column schematic

control jets can be located at positions 1 to 4 along the mast, and reaction control jets can be located at positions around the hoop. The antenna characteristics are shown in Table 2.

Table 2. Antenna Characteristics

Diameter = 400 ft.	MODE	ω_n , rad/sec	Period, sec
$W = 10020 \text{ lb}$	1	.7466	8.42
$I_x = 1.360 \times 10^8 \text{ lb-ft}^2$	2	1.346	4.67
$I_y = 1.365 \times 10^8 \text{ lb-ft}^2$	3	1.7025	3.69
$I_z = 1.041 \times 10^8 \text{ lb-ft}^2$	4	3.1813	1.98
$I_{xz} = 0.58 \times 10^6 \text{ lb-ft}^2$	5	4.5294	1.39
	6	5.5905	1.12

Various control techniques are being investigated to establish the most reliable and efficient design procedures for the hoop-column antenna. Figure 16 illustrates typical results for two of the design methods investigated. The Linear Quadratic Gaussian (LQG) optimal control technique is being used as the basis for comparison of the various techniques. The plots show that the LQG and decoupled control methods produce comparable results for nulling disturbances in the three rigid-body modes and in the first three flexible (vibration) modes. Of the two methods, the LQG method is the more complicated because of the iterative process required to achieve desired performance. The decoupled method provides a simple closed-form solution and produces the exact closed-loop dynamics

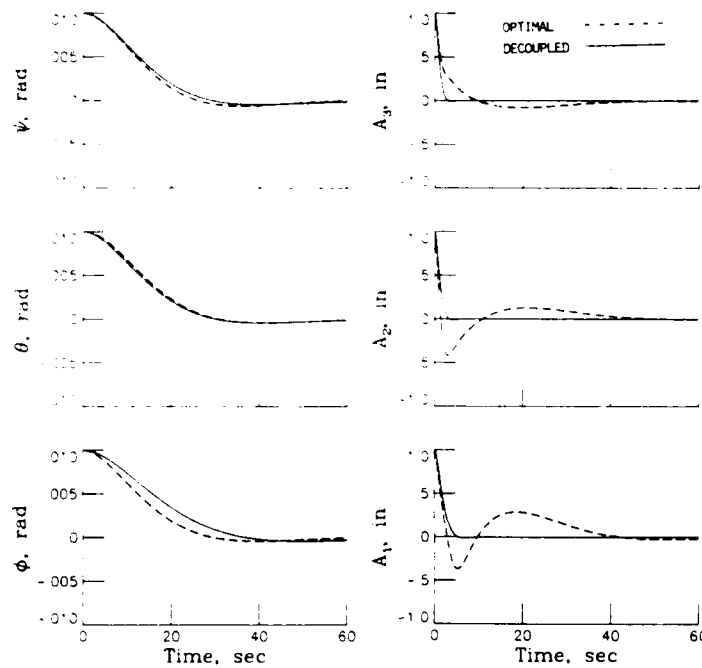


Figure 16. Comparison of optimal and decoupled responses

specified for the system. In another investigation control systems were designed using colocated output feedback and LQG techniques. Figure 17 shows a comparison of performances obtained using the two methods. The controllers were designed for three different closed-loop, rigid-body bandwidths (ω_B), and for various values of the (closed-loop) time constant corresponding to the structural modes. The horizontal axis for the colocated controller figures indicates the desired negative shift in the closed-loop eigenvalues corresponding to the flexible modes, which is used as a design parameter. The horizontal axis for the LQG controller indicates successive increase at each step in the "Q" matrix, which indirectly achieves the same effect. The expected values of pointing errors, feed-motion errors, and surface errors were computed in the presence of sensor and actuator noise. Although both controllers satisfy the requirements, the LQR controller is found to perform significantly better, especially when a smaller number of sensors and actuators was used.

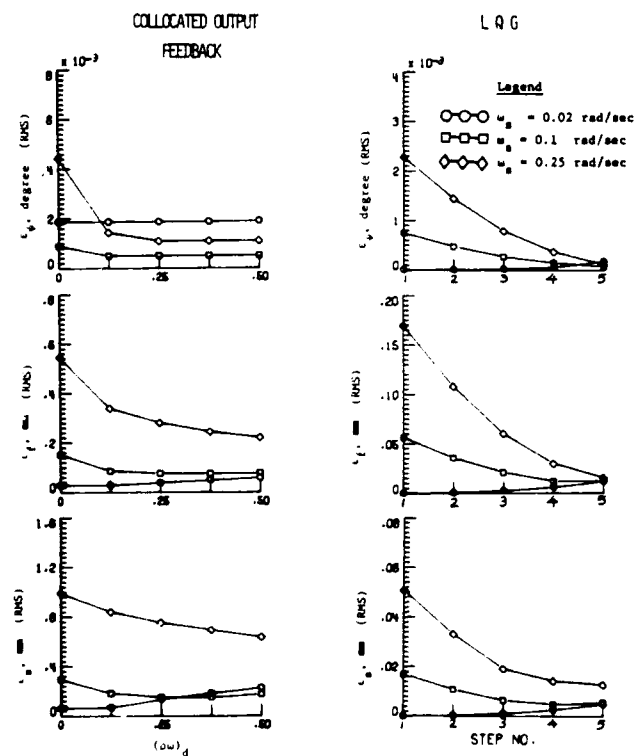


Figure 17. Controller performance analysis

Large deployable reflector

The second major class of large antenna missions corresponds to applications where the wavelength is so short that solid reflectors are required. Typical of this class is the Large Deployable Reflector (LDR) shown in Figure 18, which is an Earth-orbiting astronomical observatory operating from 50 to 1000 μm , a region of the spectrum where ground-based mapping of the sky is severely limited due to atmospheric opacity. For the present LDR concept shown the segments are deployed from their folded configuration into final reflector shape. To satisfy the observational requirements, the position and orientation of each segment must be sensed and controlled to extremely high precision. The major components are the segmented primary reflector, the secondary reflector, the backup structure, the spacecraft bus with its cryogenically cooled focal plane instruments, the solar arrays, and the thermal baffle.

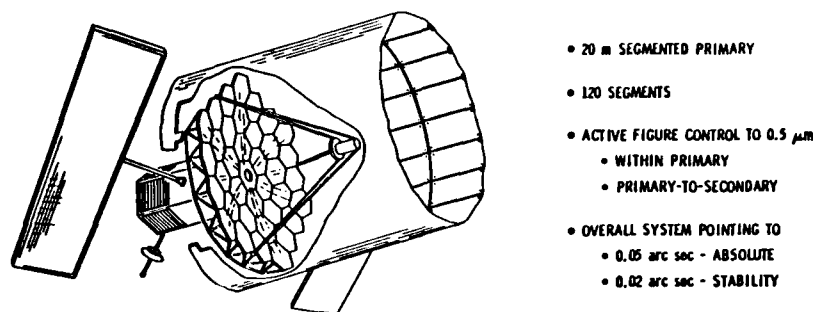


Figure 18. Large deployable reflector

The potential control approach shown in Figure 19 calls for an attitude control system with wheels and thrusters to steer and stabilize the focal-plane assembly, with attitude sensing provided by the Guide Telescope and LOS (line of sight) transfer system. This ultra-precise IR LOS pointing of the focal plane is accomplished by transmitting a laser beam from the guide telescope to the focal plane via a set of mirrors, a selected segment reflector surface, and the secondary reflector. Fine pointing at the focal plane can be enhanced by means of a fast-steering mirror and detector electronics. The direction of the laser beam is determined by the guide telescope, which uses a star tracker, IRU (Inertial Rate Unit), and its own attitude controller to guide its orientation relative to the stars.

The shape of the primary reflectors is maintained by the Primary Figure Controller, which drives the segmented reflectors against the backup structure. The primary figure sensor unit is mounted in the vicinity of the secondary reflector, which is driven by a suitable drive mechanism actuating against its support structure.

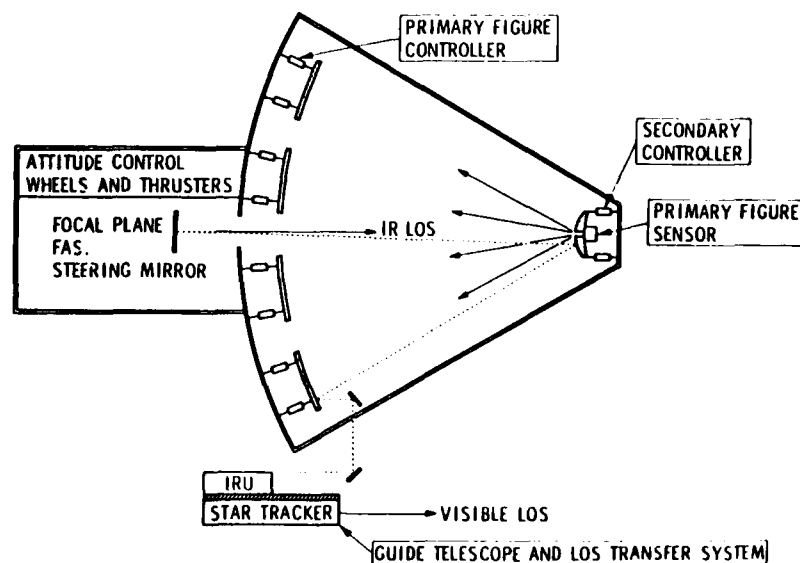


Figure 19. LDR control approach

COMPONENT DEVELOPMENT

Within the NASA control and guidance program we are also pursuing the development of sensors and actuator components to support LSS. Shown in Figure 20 is a shape determination system (SHADES) which is based on the use of a multipoint optical sensor.

SHADES

- AN INTEGRATED HARDWARE/ SOFTWARE SYSTEM FOR SHAPE DETERMINATION AND SYSTEM IDENTIFICATION OF LARGE FLEXIBLE STRUCTURES

FEATURES

- SELF-CONTAINED PORTABLE MODULE (SUITCASE)
- MULTIPOINT OPTICAL SENSING VIA SPATIAL HIGH-ACCURACY POSITION ENCODING SENSOR (SHAPES)
- RECURSIVE MAX LIKELIHOOD ESTIMATION ALGORITHMS FOR ON-BOARD IMPLEMENTATION
- INTEGRATED SHAPE ESTIMATION / IDENTIFICATION

APPLICATIONS

- LSS FLIGHT EXPERIMENT
- GROUND/SPACE ANTENNAS

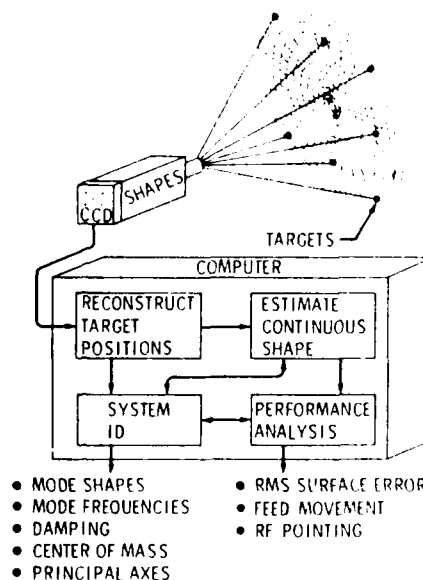
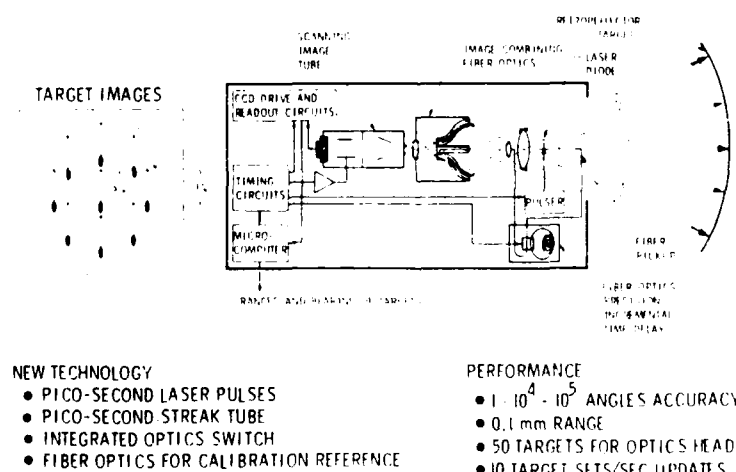


Figure 20. Shape determination system

The aspects of reduced data requirements and accurate autonomous sensing are the key features of SHADES. In addition to the shape estimator design, SHADES will contain multipurpose subroutines performing such functions as system identification, tracking of feed movement, and RF pointing. SHADES is an integrated system and is made up of two fundamental technologies: multipoint spatially distributed sensing, provided by SHAPES (Spatial High-Accuracy Position Encoding Sensor), and estimation and identification methodology for processing the sensor data in order to establish knowledge of the vehicle's static and dynamic shape. In addition, SHADES will contain the system identification algorithms required for estimating poorly known parameters (modal frequencies, damping, etc.) in the dynamical models. While the figure illustrates an application of SHADES to a large antenna, the basic sensing, estimation, and identification capabilities of the integrated system are more generally applicable to a wider range (platforms and space stations) of large structure configurations.

The SHAPES multipoint sensor concept shown in Figure 21 is designed to meet the need for a shape determination and system identification sensor that can measure the three-dimensional coordinates of many points on a large space structure simultaneously. It utilizes recent developments in short-pulse laser diodes, fiber optics, integrated



NEW TECHNOLOGY

- PICO-SECOND LASER PULSES
- PICO-SECOND STREAK TUBE
- INTEGRATED OPTICS SWITCH
- FIBER OPTICS FOR CALIBRATION REFERENCE

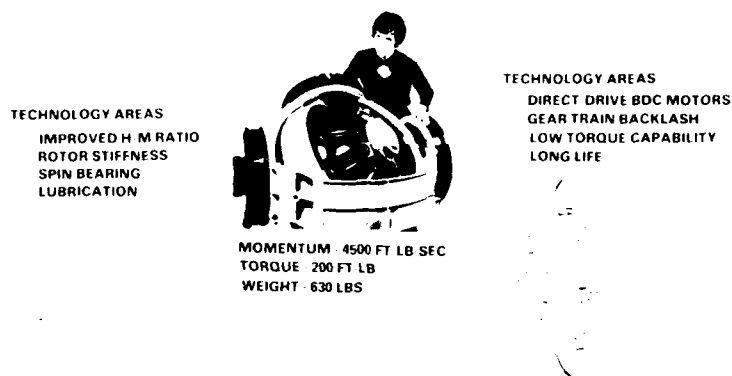
PERFORMANCE

- 10^{-4} - 10^{-5} ANGLES ACCURACY
- 0.1 mm RANGE
- 50 TARGETS FOR OPTICS HEAD
- 10 TARGET SETS/SEC UPDATES

Figure 21. SHAPES sensor

optics, streak tubes, and CCD technology. It operates by measuring the range to and angular position of targets on the structure from a single location. Range is measured by the time of flight of very short light pulses. The precision of this measurement is improved by the use of a discrete-delay reference path consisting of fiber-optics links switched by integrated-optics switches. The time delay is determined by a streak tube with a CCD readout. The angular position is determined by the direct imaging of the targets on a second CCD. The projected performance for SHAPES is 10 readings per second from up to 50 targets per optics head with range uncertainties of 0.15 mm and angular position uncertainties of 10^{-4} of the field of view. More targets can be accommodated in parallel with the addition of a second optics head.

In the control actuator area for large structures we are looking once again at large control moment gyros (CMG) for primary control torque and momentum management. The largest CMG ever developed is shown in Figure 22 which has a maximum momentum capacity of 4500 ft-lb-sec. This CMG was actually developed in the early to mid-1970's when earlier space station concepts were seriously under study. After reaching a relatively advanced state of development the spin assembly was tested with the gimbals locked and without the torquers. The test results at that time looked encouraging. In the current program the assembly has been taken out of storage, and once again evaluated. The test data appear to indicate that this large CMG technology is at a technology readiness level and the torquer represents an efficient and well designed actuator ready for integration with the wheel assembly. For such a CMG used in the control of a large space structure a number of benefits become apparent. Use of such a large CMG implies fewer desaturation maneuvers, larger momentum infers fewer units for control and thus lower weight/volume, and fewer units result in reduced software complexity.



APPLICATIONS: MANNED SPACE STATION, LARGE SPACE STRUCTURES, LONG DURATION SATELLITES

BENEFIT:

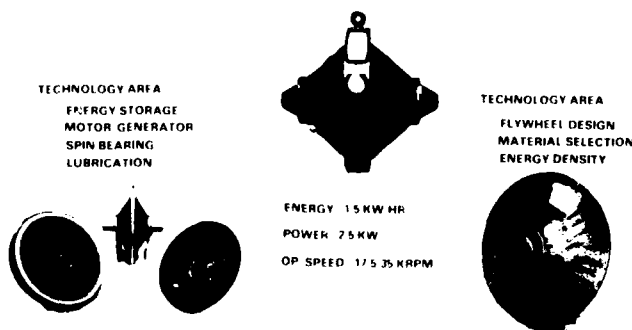
LARGER MOMENTUM PERMITS FEWER DESATURATION MANEUVERS

LARGER MOMENTUM MEANS FEWER UNITS, THUS LOWER WEIGHT/VOLUME

FEWER UNITS RESULTS IN REDUCED SOFTWARE COMPLEXITY

Figure 22. Control momentum gyro (CMG)

The effectiveness of momentum storage devices for providing control of a space vehicle's attitude has been demonstrated in flight. However, in most cases, the full potential of these devices was not utilized. One approach to realizing the full capability of a gimballed momentum device is to combine the control functions associated with such a unit with those attendant with energy storage and power generation. The unit shown in Figure 23 is the rotating assembly of a double-gimbal IPACS (Integrated Power/Attitude Control System) unit. This device is capable of storing 1.5 kW-hr of energy and of delivering 2.5 kW of power at 52 Vdc. In addition, it is designed to provide 20 ft-lb of torque required for control of an unmanned spacecraft. The research on this device will concentrate on establishing the generic technology associated with such a concept to permit future applications. These efforts will consist of characterization testing, math modeling, and utilization of such models in system trade-off analyses.



APPLICATIONS: SPACE STATION, LARGE SPACE STRUCTURES, LONG DURATION SATELLITES

BENEFIT:

REPLACES BATTERIES RESULTING IN WEIGHT/VOLUME SAVINGS

LONG-LIFE OPERATION

Figure 23. Integrated power/attitude control system (IPACS)

KEY TECHNOLOGY NEEDS

Driver missions which have not been mentioned, outside of the LSS area, which also pose advanced control technology requirements, are those in the advanced earth observation spacecraft area, such as the next generation LANDSAT spacecraft, and the Space Telescope. In addition we have activities in advanced technology development for advanced shuttle in order to make it more cost affordable. Looking at all of the control requirements driven by future missions over the next decade, we see there are a host of key technology needs before us as shown in Figure 24. Implicit in this list of needs for control of large structure is the area of fault tolerance research. The control and stabilization system for large space structures is likely to have upwards of several hundreds of components distributed over the structure and linked together by some form of data distribution networks. Even with extremely reliable components, there are predictably going to be numerous component failures or equivalent hardware failures through software faults over the lifetime of the spacecraft. The system must have the capability of accommodating these faults through system performance and fault monitoring, failure detection and isolation (FDI), and fault tolerant processing and data distribution and control law reconfiguration. Therefore, the overall problem of designing a control system for fault tolerance must be given attention in our future development programs. The last item in Figure 24, ground and flight validation is an area of extreme importance because the acceptance and commitment for use of new technology products may follow only after "technology readiness" has been established. Verification and validation of the technology through ground and flight test programs become essential. However, in the control of flexible structure the control technology cannot be validated on its own because of the close coupling and interaction with structure. This leads us to another key technology area that needs to be addressed, the technology of control and structure interaction.

CONTROL/STRUCTURE INTERACTION

The technology for controlling the attitude and dynamic deformations of large space structures is one of the key considerations for future space initiatives. Active control of flexible structures (ACOFs) technology deals with the active suppression of flexible body responses, as distinct from present practice of control of rigid body motions and avoidance of flexible structure and control interaction.

Major problems

The major problems which drive interaction between control and structure are shown in Figure 25. It is clear that the control problems represented here are many times more complex than the control problem experienced in previous generations of satellites. In general the control system must be able to provide two or three orders of magnitude of vibration alteration beyond what is obtained through natural damping or by means of viscoelastic surface treatment. To do so, it may be necessary to have as many as 100 flexible modes in the control bandwidth, although only some of these modes will be excited by the disturbances, and fewer yet will contribute to the control cost function. It should be noted that the problems listed are typically associated with large lightweight, flexible structures, but they may in fact be common to smaller precision spacecraft where performance is not achievable without accommodation for these drivers on the technology.

- | | |
|--|--|
| • MODELING AND MODEL REDUCTION TECHNIQUES | |
| • CONTROL AND STABILIZATION OF FLEXIBLE STRUCTURES | • LOW FREQUENCY AND DENSE CLOSELY COUPLED MODES |
| MODAL CONTROL | ACTIVE VIBRATION DAMPING |
| DISTRIBUTED SENSING & ACTUATION | DISTURBANCE ISOLATION |
| DECENTRALIZED CONTROL | FIGURE CONTROL |
| MODULAR, EVOLUTIONARY CONTROL | RENDEZVOUS & DOCKING |
| • UNCERTAINTY MANAGEMENT | • UNCONTROLLED DEGREES OF FREEDOM (SPILLOVER) |
| SYSTEM IDENTIFICATION | • UNKNOWN STRUCTURAL DAMPING |
| ADAPTIVE APPROACHES | • DEPLOYMENT DYNAMICS AND RELIABILITY |
| • HARDWARE AND SOFTWARE | • IDENTIFICATION OF SYSTEM PARAMETERS |
| • GROUND AND FLIGHT VALIDATION | • DISTORTION, DEFLECTION AND RECOVERY REQUIREMENTS |
| | • NONLINEAR DYNAMIC BEHAVIOR |

Figure 24. Key technology needs

Figure 25. Major problems for control/structure interaction

Ground based experiments

Over the past several years NASA's controls research program has evolved many theoretical concepts to address deficiencies cited earlier for control of large flexible spacecraft. The program has also involved experimental activities to guide and validate the theoretical work. The experimental work employs ground experiments of a generic nature with test articles comprised of beams, grids, space trusses, etc., which are likely candidates for spacecraft application. Flexible beam facilities have employed vertically suspended beams, hinged at the end, and horizontal beams suspended by wires. The facility has been augmented with detail finite-element models of the flexible beam, accurate calibrations of the sensors and actuators which were distributed along the beam, a highly interactive software package for implementing various control systems, and laser hardware with beam-mounted retroreflectors for vivid display of the beam's motion. Also, as could

only be found with actual hardware, nonlinearities, static friction, hysteresis, and unmodeled modes severely altered the control-system design process. The beam facilities have provided useful hands-on experience to researchers within NASA and to university and industry investigators in carrying out experimental tests of new control concepts. Principal experiments have involved shape estimation and control, distributed control systems, and model adaptive and insensitive control approaches. A major thrust of future work will involve development and test of advanced adaptive techniques, distributed control, and shape and vibration control. In past years static shape control and vibration control have been independently demonstrated. Future work will be aimed at combining these distinct modes of operation. Control of distributed parameter systems based on continuum models will be investigated further to allow for generalized sensors (rate, acceleration, angular, strain, etc.) and possibly generalized actuators. Shape control will be performed on more complex, multidimensional structures such as plate-like and grid-type structures.

Another ground based control demonstration deals with the problem of stabilizing and pointing flexible bodies mounted to a base structure. The test setup is shown in Figure 26. The test article is a Voyager magnetometer boom approximately 11 m x 18 cm, and the control system is built around a two-axis gimbal system developed for the Advanced Gimbal System (AGS) program. Sensors mounted on the gimbal plate and also at the far end of the boom are connected to a microprocessor containing the control logic. The structure is vertical with both the upper suspension and the base supported on air bearings. Tests will be conducted to investigate robust controller designs with structural frequencies within the control bandwidth.

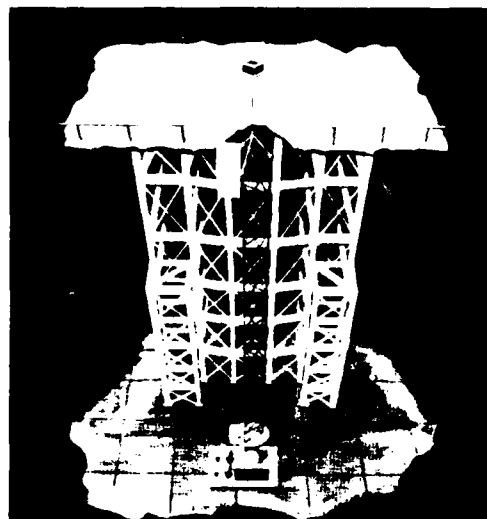


Figure 26. Boom control experiment

Flight experiments

As mentioned earlier, validation of technology through ground and flight test programs may be essential to the acceptance and application of new technology products. The structures and controls community has acknowledged through NASA and DoD workshops that uncertainties with large structure ground tests will be too great and the results cannot be relied on. Although we expect future ground testing techniques and facilities to be advanced considerably there will still exist a serious deficiency to validate control and structure technology in one "g" with equivalent full scale models. Therefore, flight technology experiments will eventually be required to demonstrate and validate technology for control of large flexible structures. Definition and design for such experiments has been underway for the past year with a focus on two types of flight experiments, MAST and a large antenna, which have control/structure experiment objectives. MAST is a Shuttle-attached multi-flight experimental research program conducted in the laboratory of space. MAST is a multi-discipline effort in structures/structural dynamics and controls employing boom type structures with configurational alternatives as shown in Figure 27. With the addition of distributed actuators and sensors and a flight controls computer, evaluation of multivariable control techniques can be initiated using the baseline cantilevered mast. With the addition of structural appendages with selected characteristics, the closely spaced modal frequencies and complex coupled motions required for the LSS research program can be produced. These phenomena model configurations will be designed to achieve the full range of characteristics necessary for a broad flight experiment program in LSS controls and dynamics. The first flight of MAST is being considered for 1987-1988.

Another technology program under consideration with a possible flight test in the 1989 time period involves a large deployable antenna in the 50 to 100 m range which also may be flown as a Shuttle-attached experiment. A need exists for such a flight experiment because of the many uncertainties with large antennas. These uncertainties result mainly from the fact that the size will increase 5 to 10⁴ over current systems and major issues cannot be resolved in ground tests and simulations. Realistic performance testing is not possible in 1 "g" and analysis with uncertainties cannot substitute for testing. A principal objective for the large space antenna flight experiment will be to demonstrate and validate a process for the design, test, and operation of large space systems with control/structure dynamic interaction.

CONFIGURATIONAL ALTERATIONS

- STRUCTURAL APPENDAGES
- DISTRIBUTED ACTUATORS SENSORS
- FLIGHT CONTROL COMPUTER

EXTENDED RESEARCH CAPABILITY

- CLOSELY SPACED MODAL FREQUENCIES
- MODES WITH COMPLEX COUPLED MOTION
- MULTI-VARIABLE CONTROL
- ON-LINE SYSTEMS IDENTIFICATION

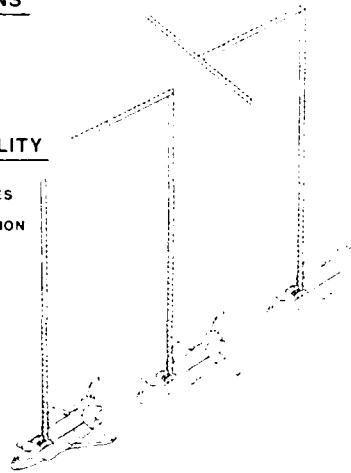


Figure 27. MAST baseline expansion

Control system definition

For the large antenna flight experiment the control system definition will involve the tasks shown in Figure 28. It is important that the experiment objective and approach address the high priority issues shown. It is likewise crucially important that results from the experiment correlate with the data from the ground test program. Although the technology program described with a large antenna flight experiment is not expected to enable technology for all large space systems it is felt that this experiment is necessary to add to the validation of the majority of technologies associated with control/structure interaction.

CONTROL EXPERIMENT OBJECTIVES AND APPROACH

DISTRIBUTED CONTROL
SYSTEM ID
ADAPTIVE/MODULAR CONTROL
STRUCTURAL DESIGN CRITERIA
GROUND TEST DATA CORRELATION

EXPERIMENT MOUNTING TRADES

RIGID
HINGES
GIMBAL MOUNTS

EXPERIMENT TEST SEQUENCES AND TIMELINE

GROUND TEST PROGRAM
PRELIMINARY CONTROL EXPERIMENT FLIGHT ASSEMBLY (EFA) CONFIGURATION

Figure 28. Control system definition

CONCLUSIONS

Future missions in space will require controlling spacecraft which are both large and flexible. The limited inherent damping and the uncertain and changing dynamic characteristics of many of these vehicles, such as manned space stations and large antennas, will revolutionize spacecraft control requirements. For applications requiring high system performance, control synthesis techniques must address model spillover and be robust and sometimes adaptive. Initial on-orbit testing will be used to more accurately model the dynamics of a new large space structure, so that its control law can be fine tuned to enhance system performance and to ensure stability.

In preparation for the time that such control systems are required, considerable research and technology development is necessary. A program is in place at NASA for the development of active control technology to support major initiatives for space station and advanced spacecraft. Space station is characterized by requirements for evolutionary and modular growth presenting new challenges in distributed modular control, utilizing forms of systems identification for multi-state active control. Similar requirements press the technology needs for spacecraft systems such as large communication satellites and submillimeter wave astronomy missions. A number of key control technology needs are required for these and other future NASA missions together with an integrated control/structures technology flight experiment to demonstrate and validate technology for large flexible structures.

This paper reviewed the general requirements of controlling large, flexible spacecraft and outlined related NASA research activities. Future opportunities with the shuttle will lead to large space systems with increased flexibility and unprecedented interaction between control and other subsystems. New control theories, sensors, actuators must be developed to achieve and maintain the required performance. Much work will be necessary in understanding integration issues of such technology through ground and flight experiments.

REFERENCES

1. Taylor, L. W., Jr., and Lin, J. B., The Challenge of Controlling Large Space Systems, IEEE Large Scale Systems Conference, Virginia Beach, VA., Oct. 11-13, 1982.
2. Dahlgren, J. B., Active Control Technology Development at NASA, SPIE 1983 Technical Symposium, Session on Structural Control of Deployed Optical Systems, Paper number 363-21, Jan 1983.
3. Montgomery, R. C., Structural Control Research and Experiments at NASA Langley, SPIE 1983 Technical Symposium, Session on Structural Control of Deployed Optical Systems, Jan 1983.
4. Taylor, L. W., Jr., and Balakrishnan, A. V., A Laboratory Experiment Used to Evaluate Control Laws for Flexible Spacecraft--A NASA/IEEE Challenge, 4th VPI & SU/AIAA Symposium on Dynamics and Control of Large Structures, Blacksburg, VA, June 6-8, 1983.
5. Williams, J. P., and Montgomery, R. C., Experimental Implementation of Parameter Adaptive Control on a Free-Free Beam, 4th VPI & SU/AIAA Symposium on Dynamics and Control of Large Structures, Blacksburg, VA, June 6-8, 1983.
6. Montgomery, R. C., Reliability Considerations in the Placement of Control System Components, AIAA Guidance and Control Conference, Gatlinburg, TN, Aug 15-17, 1983.
7. Sundararajan, N., and Williams, J. P., Parameter Testing for Lattice Filter Based Adaptive Modal Control Systems, AIAA Guidance and Control Conference, Gatlinburg, TN, Aug 15-17, 1983.
8. Garibotti, J. F., et al., Final Report - Space Systems Technology Advisory Committee Ad Hoc Subcommittee on Controls/Structure Interaction, NASA Report, June 8, 1983.
9. Cameron, J. M., et al., Static Shape Determination for Large Structures, 4th VPI & SU/AIAA Symposium on Dynamics and Control of Large Structures, Blacksburg, VA, June 6-8, 1983.
10. Rodriguez, G., Hamidi, M., and Eldred, D., Maximum Likelihood Identification of Parameters in Static and Dynamic Shape Estimator Models, 4th VPI & SU/AIAA Symposium on Dynamics and Control of Large Structures, Blacksburg, VA, June 6-8, 1983.
11. Wang, S. J., Lin, Y. H., Ih, C. C., Dynamics and Control of a Shuttle Attached Antenna Experiment, AIAA Guidance and Control Conference, Gatlinburg, TN, Aug 15-17, 1983.
12. Tolivar, A. F., and Wang, S. J., Control of Large Space Antennas, NASA Conference on Large Space Antenna Systems Technology, Nov 1982.
13. Keckler, C.R., et al., Description and Simulation of an Integrated Power and Attitude Control System Concept, NASA TN D-7459, Apr 1974.
14. Buchanan, H. J., Practical Approaches to the Design of Control Systems for Large Space Structures, NASA Conference on Large Space Antenna Systems Technology, Nov 1982.
15. Montgomery, R. C., and Taylor, L. W., Jr., Results of Studies at Langley Research Center on the Control of Large Space Systems, NASA Conference on Large Space Antenna Systems Technology, Nov 1982.

ABBREVIATIONS USED IN TEXT

R&T	Research and Technology
NASA	National Aeronautics and Space Administration
LSS	Large Space Systems
ID	Identification
RF	Radio Frequency
LMSS	Land Mobile Satellite Service
LQG	Linear Quadratic Gaussian
IDR	Large Deployable Reflector
SHADES	Shape Determination System
SHAPES	Spatial High Accuracy Position Encoding Sensor
CMG	Control Moment Gyro
IPACS	Integrated Power/Attitude Control System
ACOFs	Active Control of Flexible Structure
MAST	Multi-flight Experimental Research Program

ROLF DE L'INTELLIGENCE DE BORD ET DE L'ENTRETIEN EN ORBITE DANS L'ACCROISSEMENT DE LA DUREE DE VIE DES SATELLITES

Par C. COUGNET, J.P. SOTTA

MATRA ESPACE
1, rue des Cosmonautes
Zone Industrielle du Palays
31400 TOULOUSE
FRANCE

RESUME

L'accroissement continu de la durée de vie des satellites est ou va bientôt devenir critique vis-à-vis du maintien de la fiabilité du système et du risque d'obsolescence de la mission.

Le recours à la redondance des équipements pour assurer la fiabilité devient une solution de plus en plus pénalisante compte-tenu en particulier des capacités limitées des moyens de lancement. D'autre part, la conception classique des satellites ne permet pas d'éviter l'obsolescence de la mission sur de longues durées de vie.

Pour remédier à ces problèmes, deux méthodes possibles complémentaires sont présentées :

- L'utilisation de l'intelligence à bord des satellites, permettant par une meilleure gestion de la configuration et des sous-systèmes de la plate-forme, d'assurer un haut niveau de disponibilité au satellite.
- L'entretien en orbite, permettant de réajuster périodiquement la fiabilité et l'autonomie de la plate-forme et autorisant également l'éventuel échange de la charge utile. Cette solution fait appel à des techniques nouvelles : le rendez-vous, l'assemblage et la robotique spatiale.

12.1 - INTRODUCTION

La durée de vie opérationnelle des satellites actuellement en orbite est de l'ordre de 7 à 10 ans pour les satellites de télécommunications et de 3 à 5 ans pour ceux situés en orbite basse. Si certaines missions s'accommodent assez bien de cette limitation, voire génèrent elles-mêmes des exigences inférieures, une grande partie bénéficieraient d'une plate-forme de durée de vie accrue, tant d'un point de vue service que d'un point de vue économique : ce pourrait être le cas de missions de télécommunications, de navigation ou d'observation.

De fait, l'analyse de la durée de vie spécifiée des satellites de télécommunications montre que celle-ci, qui était de l'ordre de 1 à 3 ans au début des années 60, est passée successivement à 5 ans, puis 7 ans (NATO III, INTELSAT IV, ...). Actuellement, les satellites de nouvelle génération sont conçus pour des durées de vie de 7 à 10 ans (INTELSAT 6).

Toutefois, l'accroissement de la durée de vie des satellites ou des plates-formes, ne va pas sans contrepartie ; ainsi par exemple, une augmentation corrélative de leur masse qui risque de se heurter aux capacités des lanceurs, l'utilisation d'équipements compatibles avec cette durée de vie, l'assurance d'un certain niveau de fiabilité et de disponibilité pendant toute cette durée.

12.2 - CAUSES DE LIMITATION

Concevoir une plate-forme pour une certaine durée de vie, c'est concevoir chaque sous-système, chaque ensemble d'équipements pour cette exigence. Cela suppose également que toutes les causes possibles de limitation de la durée de vie ont été prises en compte de sorte qu'elles n'apparaissent pas avant la date prévue.

Ainsi, cinq types de limitations peuvent être considérés :

- L'épuisement des ressources non renouvelables : cet appauvrissement concerne essentiellement les ergols (propulsion et contrôle d'attitude), mais peut éventuellement toucher aussi les lubrifiants. Cet épuisement d'ergol est une cause de dégradation déterministe et peut être prédit en utilisant un modèle de consommation, si bien que la masse d'ergols nécessaire pour une durée de vie donnée peut être évaluée avec une bonne précision. Toutefois, le problème devient plus critique si des opérations coûteuses en ergols et non prévues deviennent nécessaires, car elles pourraient conduire à un appauvrissement prématuré du satellite en ergols.
- La décroissance des performances du système : cette décroissance résulte de phénomènes d'usure ou de dégradations de performances de sous-systèmes, et peut être déterministe ou statistique. Cette cause de limitation affecte principalement les panneaux solaires, les batteries, les optiques, les assemblages mécaniques, les radiateurs thermiques. Cela conduit à dimensionner certains équipements (panneaux solaires) en fonction des spécifications en fin de vie, ce qui peut résulter en un surdimensionnement en début de vie. D'autres équipements peuvent présenter des limitations intrinsèques de durée de vie du fait de leur conception ou leur technologie et deviennent alors critiques.
- Des pannes du système, qui sont essentiellement aléatoires et peuvent intervenir dans n'importe quel sous-système. Ces pannes peuvent être définitives, ou conduire à une dégradation de la mission, ou ne pas affecter la mission mais réduire la fiabilité du système.
- L'erreur humaine ou une conception inadéquate de certains éléments.
- L'obsolescence de la mission, qui est essentiellement relative à la charge utile.

Parmi ces causes de limitation, la dégradation des performances et certaines défaillances dues à l'usure pourraient être prédictibles et donc influencer le dimensionnement des équipements (panneau solaire) ou la configuration et l'utilisation de sous-systèmes (schéma de redondance, plage de fonctionnement). Les pannes aléatoires, quant à elles, peuvent être représentées par une courbe de fiabilité qui donne la probabilité de bon fonctionnement du système en fin de vie. Le schéma de redondance des équipements sera conçu pour optimiser cette probabilité.

L'apparition et l'évolution des dégradations ou des défaillances des équipements dépendent du choix de la technologie, mais elles sont également favorisées par l'environnement dans lequel se trouve le satellite : par exemple, les niveaux de radiations auxquels sont soumis les équipements (cellules solaires, revêtements thermiques, ...) et les composants sont un facteur important de dégradation et varient suivant la position orbitale du satellite (altitude, inclinaison) ; de même pour les cycles thermiques ou les perturbations diverses. Elles sont également liées aux conditions d'utilisation opérationnelle des équipements (température, cycles de fonctionnement, ...).

L'accroissement de la durée de vie d'un satellite conduit naturellement pour une technologie donnée, à un accroissement de sa masse, voire de son volume, essentiellement à cause de l'augmentation induite d'ergols et d'équipements redondants. Or, une conception harmonieuse suppose une répartition optimale entre le nombre d'équipements en redondance et la masse disponible d'ergols, de telle sorte que l'épuisement des ergols et l'apparition de phénomènes d'usure ou de dégradation du système coïncident avec la limite de durée de vie spécifiée. Une durée de vie spécifiée devrait donc correspondre à une répartition optimale des masses du satellite, cette répartition étant elle-même fonction de l'état de l'art de la technologie ; or cette répartition n'est jamais libre, ne serait-ce que parce que la masse totale du satellite est limitée par les capacités du lanceur. Cette limitation, due au lanceur, pourrait toutefois être surmontée par l'utilisation de techniques de rendez-vous et assemblage en orbite qui permettraient d'implanter des satellites ou plates-formes de masse ou volume au-delà des performances du lanceur.

La courbe de fiabilité est également un paramètre important pour la conception du satellite car elle influence la durée de vie prédictible. Cette durée de vie est fonction de la fiabilité acceptable en fin de vie, si bien que la courbe doit être ajustée à la durée de vie requise.

12.3 - SOLUTIONS ENVISAGEABLES

Chaque élément a donc sa durée de vie propre et peut devenir critique vis-à-vis d'une augmentation de la durée de vie du satellite, suivant le niveau de l'exigence.

Plusieurs solutions sont alors envisageables pour y remédier (voir fig. 1).

Une solution classique est l'auto-maintenance. Elle consiste à redondier les équipements (ou les fonctions) critiques afin de diminuer les risques de panne du système ou d'accroître la durée de vie de l'ensemble redondé au-delà des dates d'apparition possible des phénomènes d'usure de l'équipement en fonctionnement. L'utilisation de l'équipement redondant est une solution pour compenser la défaillance d'un équipement nominal et ainsi continuer la mission ; cependant, des dégradations peuvent apparaître également sur des équipements en attente, si bien que leur durée de vie, quoique supérieure à celle de l'équipement en fonctionnement, pourrait être elle-même une limitation à la durée de vie du satellite. De plus, cette solution conduit à un accroissement de la masse du satellite.

Une autre solution consiste à améliorer la technologie des équipements critiques ou à utiliser de nouvelles technologies dans le but de diminuer leur taux de défaillance ou d'augmenter leur durée de vie s'ils sont sujets à dégradation et usure, ou éventuellement d'améliorer leur rendement afin de diminuer leur masse au lancement à performance égale. C'est le cas par exemple des batteries NiH₂ dont on espère une durée de vie meilleure que celle des batteries NiCd.

Toutefois, les améliorations technologiques et les nouvelles technologies peuvent avoir également leurs propres limitations ; aussi, une autre solution peut consister à développer l'intelligence de bord du satellite pour optimiser la gestion des sous-systèmes, pour améliorer les conditions d'utilisation et de fonctionnement des équipements critiques (gestion des batteries et de leur contrôle thermique par exemple). Les avantages et les possibilités de cette solution sont décrits dans le chapitre suivant. Cette solution permet d'améliorer la courbe de fiabilité du système ou de retarder l'apparition de dégradations ou d'autres causes de limitation de durée de vie du système.

L'entretien en orbite du satellite est une solution intéressante car elle permet d'effectuer périodiquement, dans une certaine limite, une remise à niveau du satellite ; ainsi, pourrait-on échanger les équipements critiques à une date prédictible en fonction de la technologie employée ; ainsi pourrait-on intervenir en orbite dans certains cas de pannes ou pour amener des ergols supplémentaires. Toutefois, cela suppose une conception adaptée du satellite et l'utilisation de véhicules spécifiques. Les avantages et inconvénients de cette solution sont présentés dans le chapitre 5.

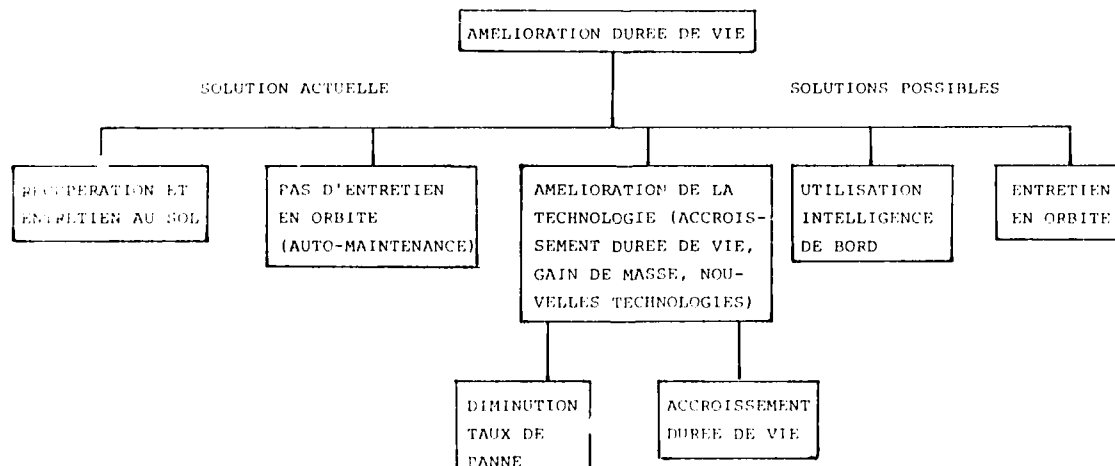


Figure 1 : SOLUTIONS POSSIBLES POUR AMELIORER LA DUREE DE VIE

12.4 - UTILISATION DE L'INTELLIGENCE DE BORD

On entend par intelligence de bord, ou intelligence embarquée, l'ensemble des boucles de contrôle automatique et des décisions prises à bord de manière autonome, sans intervention du sol en temps réel. Les satellites actuellement en orbite ne possèdent qu'un niveau rudimentaire d'intelligence de bord, qui en outre est généralement fortement spécialisée (comme le d'attitude utilisant des boucles de contrôle analogiques ou numériques, décider le passage en mode survie, etc...). Cependant, l'arrivée de micro-calculateurs compatibles avec l'environnement spatial va permettre d'augmenter l'importance et la finesse de l'intelligence embarquée dans des proportions considérables.

12.4.1 - Rôle de l'intelligence embarquée

Le rôle de l'intelligence de bord dans l'accroissement de la durée de vie des satellites peut être multiple et s'exercer au niveau système aussi bien qu'au niveau sous-système ou équipement ; en particulier, les domaines suivants peuvent bénéficier de l'utilisation de l'intelligence embarquée :

- Amélioration de la fiabilité intrinsèque de chaque fonction par l'utilisation de technologies digitales et de méthodes de contrôle par micro-calculateurs. Un seul calculateur peut gérer plusieurs boucles de contrôle en parallèle, ce qui conduit à une simplification de l'électronique de décision et à une réduction du taux de panne global car les niveaux d'intégration et de fiabilité des micro-calculateurs sont nettement supérieurs à ceux de l'électronique conventionnelle.
- Gestion plus fine des éléments consommables comme les batteries ou les ergols. La mise en oeuvre de manière automatique à bord de modèles de gestion sophistiqués permet d'optimiser en temps réel la gestion des consommables en fonction des besoins effectifs du service requis. De plus, il est possible d'introduire la notion de prévision (de charge ou de couples perturbateurs par exemple) dans ces modèles de gestion optimale.
- Amélioration de la disponibilité du service offert. Seul un haut niveau d'intelligence embarquée peut limiter les conséquences d'une panne d'un équipement et régénérer les fonctions nominales du satellite dans un délai suffisamment court.

Il reste toutefois, comme condition préalable à l'emploi de micro-électronique en orbite pendant de longues années, que la technologie sélectionnée doit présenter une bonne résistance à l'environnement spatial. Ainsi, la technologie TTL est naturellement capable d'une durée de vie en orbite d'une dizaine d'années alors que seules certaines technologies CMOS sont susceptibles de convenir, généralement grâce à des traitements appropriés.

12.4.2 - Applications potentielles de l'intelligence embarquée

L'utilisation de l'intelligence embarquée pour gérer les sous-systèmes permet d'optimiser les conditions de fonctionnement des différents équipements et, par ce biais, a une action bénéfique sur la durée de vie.

12.4.2.1 - Sous-système de génération et distribution d'énergie électrique

Pour chacune des trois fonctions principales du sous-système, à savoir production et stockage d'énergie et adaptation à la charge, l'utilisation d'intelligence embarquée conduit à une amélioration des conditions d'utilisation des éléments critiques.

- La génération de puissance est souvent effectuée au moyen de panneaux de cellules solaires qui présentent un point de rendement maximum dépendant des conditions thermiques et du vieillissement. Il est possible de maintenir le point de fonctionnement au voisinage de l'optimum en faisant varier en permanence la charge de la batterie en sens inverse de la consommation du satellite. Une telle boucle de contrôle nécessite la puissance et la souplesse d'un micro-calculateur, et elle optimise simultanément le rendement des générateurs solaires et la charge de la batterie. Le gain en durée de vie utile des panneaux solaires est d'autant plus important que le profil de consommation de la mission est variable. Dans son principe, une telle boucle de contrôle est également valable pour un générateur isotopique.
- Pour limiter le vieillissement de la batterie, les opérations de charge et decharge et les cycles d'entretien doivent respecter des procédures très précises qui, en pratique, ne peuvent être pleinement mises en oeuvre sans un micro-calculateur embarqué. En effet, la gestion de la charge de la batterie doit être couplée avec l'optimisation du point de fonctionnement des générateurs solaires et les variations de consommation électrique du satellite et doit tenir compte des prévisions de consommation à moyen terme (quelques révolutions, par exemple).

De plus, ce modèle de gestion doit intégrer la notion de temps de réponse à une requête sol. Il apparaît clairement que l'optimisation des rendements du système de génération de puissance suppose la prise en compte simultanée de nombreux paramètres, allant du plus simple, comme la température batterie, à des notions complexes de philosophie opérationnelle. Le rôle de l'intelligence de bord devient alors évident et nécessaire.

12.4.2.2 - Sous-système de contrôle thermique

La température moyenne de fonctionnement est un facteur important dans le calcul des taux de panne d'un équipement électronique ou mécanique, et la fiabilité résultante sera d'autant meilleure que la température sera modérée.

Le pilotage par logiciel bord permet d'exploiter au maximum les systèmes actifs de contrôle de température surtout lorsque l'intelligence de bord a la faculté de prévoir l'évolution du bilan des échanges thermiques. L'amplitude des régimes transitoires et les erreurs de trainage sont ainsi très nettement réduites, et en conséquence, le "stress thermique" des équipements bord l'est également.

Il faut noter que la prévision de l'évolution des températures nécessite :

- Un modèle dynamique simplifié des interactions radiatives entre le satellite et l'ensemble espace-soleil-terre.
- Un modèle dynamique des échanges calorifiques internes. La précision de prédiction à quelques heures de l'ordre de $\pm 5^{\circ}\text{C}$ devrait suffire pour maintenir les températures dans les plages spécifiées à $\pm 10^{\circ}\text{C}$. Si l'on admet une période de rodage en début de mission, la valeur exacte des coefficients de couplage peut être recalée par le sol en fonction des résultats des premières orbites. Il est possible aussi de prendre en compte la dégradation au cours du temps des revêtements thermiques par une remise à jour périodique d'un jeu de constantes.
- La corrélation entre le plan d'opération stocké à bord et la variation afférente des dissipations thermiques, en particulier les périodes d'utilisation de la charge utile et des opérations de charge-décharge de la batterie doivent être prédites.

Certains équipements peuvent nécessiter une régulation thermique fine, réalisée par une boucle de contrôle spécifique, distincte du système général.

La précision d'une boucle thermique par logiciel dépend essentiellement du nombre de capteurs/actuateurs et de la fréquence de répétition. Les techniques logicielles impliquées sont donc surtout de nature base de données.

12.4.2.3 - Sous-système contrôle d'attitude et d'orbite

Ce sous-système utilise traditionnellement un niveau élevé d'intelligence embarquée pour assurer la majeure partie de ses fonctions. L'accroissement de la durée de vie de ce sous-système passe essentiellement par :

- une gestion parcimonieuse des ergols,
- le report de la mise hors service définitive des capteurs d'attitude qui sont sujets au vieillissement,
- une meilleure adaptation des lois de contrôle à l'évolution des performances des couples parasites.

Dans les satellites modernes, les ergols consommables ne sont utilisés que dans les modes transitoires (desaturation des roues, circularisation d'orbite, etc...), le mode normal employant généralement des énergies renouvelables comme le moment cinétique ou la pression solaire.

Le rôle de l'intelligence de bord est alors :

- d'éviter la manoeuvre, quand c'est possible. C'est le cas si l'utilisation de l'intelligence de bord permet de limiter les conséquences d'une panne, en particulier d'éviter le passage direct en mode survie qui généralement est fortement consommateur d'ergols,
- de retarder la manoeuvre au moyen d'une prédiction de la valeur et du sens des couples perturbateurs et en ne corrigeant que les dérives à moyen terme. Ceci nécessite que, de par la conception du système, l'amplitude des perturbations périodiques n'induisse pas d'erreurs prohibitives,
- d'optimiser le rendement de chaque impulsion en contrôlant finement sa durée et sa date d'application.

L'information délivrée par les détecteurs d'attitude est généralement bruitée, d'autant plus que le détecteur vieillit (bruit thermique croissant des détecteurs infra-rouges, bruit de roulements des gyros frottés, etc...). L'utilisation de système de filtrage à micro-processeurs permet de réduire la bande de bruit et de prolonger la durée de vie effective des détecteurs. Les filtres numériques (Kalman, par exemple) permettent de pondérer différemment et de manière aisée, chaque type de détecteur d'une chaîne de contrôle, ce qui permet des rééquilibres en fin de vie. Cette méthode nécessite de tenir à jour les

caractéristiques de chaque équipement, donc d'utiliser les techniques des *lines de command*.

Les méthodes de contrôle adaptatif permettant d'optimiser en temps réel les caractéristiques d'un *mode de contrôle*. Elles nécessitent l'établissement de modèles des couples perturbateurs dont le point de fonctionnement est fixe de manière à minimiser l'erreur finale.

12.4.2.4 - Sous-système gestion des données

La tâche classique de ce sous-système est d'assurer la gestion des échanges d'information entre le satellite et le centre de contrôle au sol, celui-ci pouvant lui déléguer une partie des tâches de contrôle des opérations. Du fait de son rôle d'interface, le sous-système gestion des données a une architecture en fonction de la nature des sous-systèmes partenaires et du volume des tâches de gestion qui est dévolu au segment bord.

L'intelligence de bord peut être utilisée pour améliorer la fiabilité des tâches élémentaires :

- effectuer la modulation et le traitement des télécommandes,
- acquérir et limiter la télémétrie.

Elle peut aussi être utilisée pour optimiser l'emploi des ressources existantes. D'autres tâches peuvent également lui être déléguées, qui sont plus proches de la gestion globale du système bord. Parmi celles-ci, on peut citer :

- la gestion des ressources matérielles nécessaires pour garantir la fiabilité et la disponibilité du système, par exemple :

- la mise en place d'un système permettant de détecter une interruption du service,
- la mise en place d'un système de bord de manière à pouvoir localiser et circonscrire la panne,
- la mise en place d'un système de tri et d'entreposage,
- la mise en place d'un système de détection,
- la mise en place d'un système de manière à permettre au sol de la corriger complètement.

Les sous-systèmes peuvent également agir en divisant le service en plusieurs fonctions faiblement couplées entre elles et réparties sur l'architecture du satellite de manière à limiter la propagation des pannes entre ces fonctions. La détection des pannes doit être confiée à une fonction indépendante de celle à tester. La période de test de détection de panne est de l'ordre de la moitié de la durée maximum d'indisponibilité tolérable pour cette fonction (typiquement quelques secondes).

L'action corrective est d'autant plus simple à identifier que la liste en est plus réduite. Ceci suggère d'organiser le système de décision en éléments plus fins, contrôlant un nombre réduit de fonctions.

Le processeur de réinitialisation d'une fonction défaillante nécessite certaines précautions si son état dépend des états antérieurs, comme c'est le cas pour des calculateurs de sous-systèmes. Il faut alors restituer à cette fonction un contexte, lui conférant un état homogène et compatible avec celui de tous les autres sous-systèmes.

Ceci implique :

- la sauvegarde des programmes par duplication dans une mémoire de secours, à priori non volatile,
- la sauvegarde périodique du contexte, impliquant la définition des points de reprise compatible avec le système général.

La période entre deux points de reprise dépend de l'application considérée mais pourrait être de l'ordre d'une dizaine de secondes.

Après une reconfiguration réussie, le système est généralement prêt à traiter automatiquement toute autre panne survenant dans l'un quelconque des sous-systèmes à l'exception de celui venant d'être reconfiguré, car il n'est pas certain qu'une configuration complète de secours puisse être déterminée à bord facilement.

Cependant, pour un grand nombre de missions, on peut considérer comme négligeable la probabilité d'avoir, dans le même sous-système, deux pannes successives, c'est-à-dire séparées par un délai tel que le sol n'a pas eu le temps d'analyser la première panne et de modifier les consignes de configuration.

12.4.4 - Conséquences sur l'organisation du système bord

L'introduction d'intelligence à bord pour optimiser le fonctionnement de chaque élément et gérer le système complet, conduit à identifier des tâches nouvelles, à répartir entre le segment sol, le système bord et les sous-systèmes bord.

Le système bord regroupe les tâches communes à plusieurs sous-systèmes et qui sont déléguées à bord par le sol. Il existe toutefois un seuil au-delà duquel l'excès de délégation augmente la complexité du système bord et annule les avantages apportés par l'intelligence embarquée du point de vue durée de vie. Ce seuil dépend de la mission et du type de tâche considérée. Par exemple, l'estimation de la position orbitale peut mettre en oeuvre un système de senseurs sophistiqués ou bien se limiter à un simple modèle temporel des éclipses et de la rotation du soleil (suffisant pour les besoins des sous-systèmes thermiques et génération de puissance).

Les sous-systèmes de bord travaillent à partir des directives du système bord, qui fournit aussi les données et les ressources communes, et optimise leur travail en fonction de critères qui leur sont propres. Chaque sous-système ou sous-ensemble comporte une entité de gestion qui règle les problèmes locaux et ne soumet au système bord que ceux qui sont insolubles au niveau sous-système.

Les contraintes de conception, telles qu'elles résultent des analyses ci-dessus, sont les suivantes :

- Le nombre d'éléments en redondance est élevé, les cross-strapping sont nombreux et les redondances triples ou quadruples ne sont pas rares. De plus, ces redondances doivent pouvoir être gérées par des calculateurs de bord.

En conséquence, il est primordial de définir des moyens et règles de gestion systématiques des redondances tant au niveau du matériel (interfaces de données, d'alimentation, etc...) que du logiciel de maintenance (signalisation des pannes, procédure de sauvegarde et restitution des contextes, base de données, mise à jour permanente de la configuration). L'établissement d'exigences de testabilité peut faciliter la localisation des pannes.

- L'utilisation extensive du logiciel n'est effectivement bénéfique pour la durée de vie que si le taux de défaillance est très bas. Ceci impose des méthodes rigoureuses de production de logiciel, la réutilisation de produits déjà validés et la limitation de la taille de chaque logiciel quitte à fractionner un logiciel jugé trop volumineux et difficilement contrôlable. Malgré ces précautions, une erreur de conception peut se révéler après le lancement. La reprogrammation en vol devient une exigence fondamentale.
- Il faut distinguer deux types d'intelligence de bord : la première est spécialisée, comme par exemple un compresseur de données ou un filtre numérique ; la seconde est chargée de la gestion des interfaces des modes de fonctionnement et des redondances. Ces deux composantes sont généralement associées mais la première se rencontre surtout au niveau équipement ou sous-système, alors que la seconde relève plus du système en général.
- Le système doit être architecturé de manière à regrouper en sous-ensembles les fonctions qui sont fortement couplées, limitant ainsi la complexité des interfaces. Au sein de chaque sous-ensemble, un logiciel de gestion est chargé de résoudre les problèmes locaux de redondance et limite ainsi la propagation des pannes vers l'extérieur.
- L'emploi de redondances doit être limité au strict nécessaire, à la fois parce qu'elles sont généralement difficiles à réaliser, et parce qu'elles requièrent du matériel supplémentaire.
- Le satellite doit être équipé d'un système de prédiction des événements futurs, composé essentiellement de 2 fonctions distinctes :
 - . une prédiction de la position orbitale (pour les applications de gestion de l'énergie et du contrôle thermique une précision modeste suffit),
 - . la mémorisation et l'exécution d'un plan d'opérations détaillé, mis à jour et complété périodiquement par le sol. Ce plan doit être dupliqué à bord pour éviter de le perdre à l'occasion d'une panne.

En réponse à cette liste de contraintes, l'architecture proposée figure 2 vise à faciliter la gestion des redondances et l'implantation de logiciel à bord.

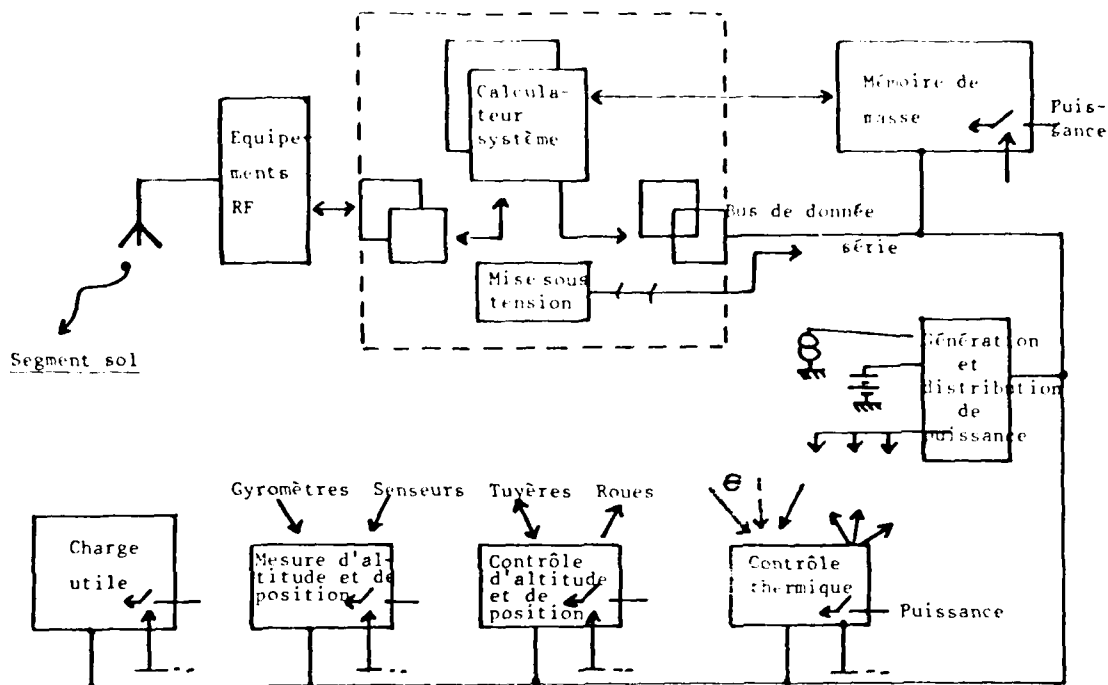


Figure 2 : ARCHITECTURE GENERALE DU SYSTEME BORD

Les équipements sont regroupés par sous-ensembles auxquels est affecté un calculateur commun chargé des tâches fonctionnelles proprement dites et de la gestion des équipements. Ces sous-ensembles sont reliés entre eux par le bus de donnée plate-forme et sont gérés par un contrôleur central. Celui-ci est en plus chargé des relations avec le sol.

Un sous-système peut regrouper plusieurs sous-ensembles communiquant à travers le bus de donnée plate-forme. Les équipements sont reliés à leur calculateur local au moyen d'un bus de second niveau dont la conception inclut non seulement l'échange de données mais aussi le contrôle des mises sous tension et des dispositifs de test. Sa conception est également fortement orientée vers le temps réel et le support à des boucles de contrôle rapides. La figure 3 donne l'organisation interne typique d'un sous-ensemble.

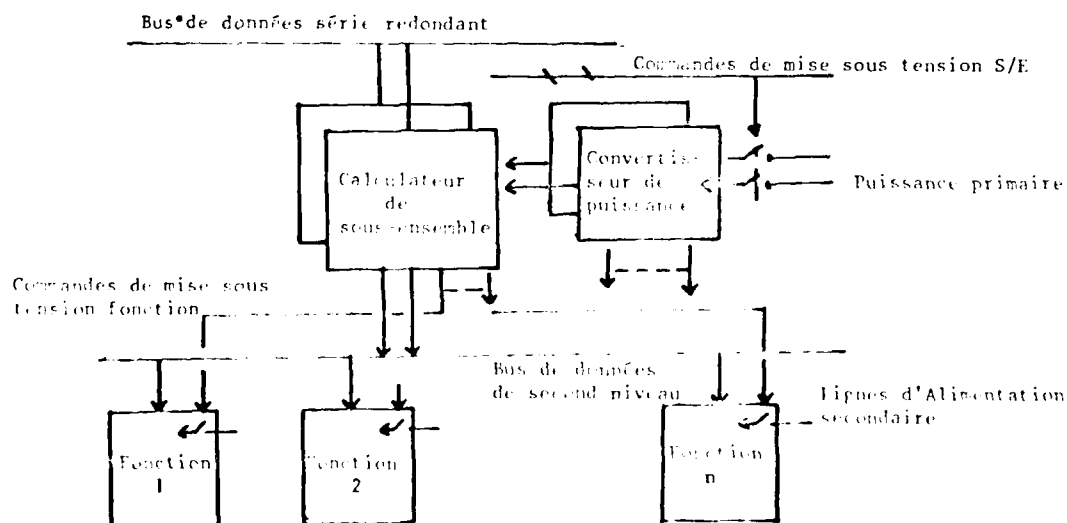


Figure 3 : ARCHITECTURE TYPIQUE D'UN SOUS-ENSEMBLE

Du point de vue de la gestion des pannes et du contrôle des modes d'opération, le système est fortement hiérarchisé. Le calculateur local, qui gère les redondances de ses propres périphériques est en redondance froide, et c'est le contrôleur central qui a la tâche de reconfigurer et de réinitialiser les calculateurs locaux en réutilisant les programmes dupliqués et les contextes sauvegardés dans la mémoire de masse. Par contre, le contrôleur central est en redondance active. Un dispositif simple, non redondé, détecte la panne du calculateur en service et en active un autre, préalablement sélectionné par le sol. Ce schéma suffit à assurer une disponibilité et une fiabilité satisfaisantes à condition que les pannes retard du dispositif de redondance soient minimisées et que la majeure partie de la mémoire de masse soit indépendante du calculateur central.

L'utilisation extensive du logiciel doit être supportée par des méthodes efficaces de validation et la possibilité de modifier en vol et par partie n'importe quel logiciel d'application. Certaines parties répétitives des logiciels système ou sous-systèmes doivent être standardisées pour en augmenter la sûreté de fonctionnement.

12.4.5 - Conclusion et limitations

L'utilisation d'intelligence embarquée apparaît donc comme un moyen intéressant pour augmenter la durée de vie des satellites.

Tout d'abord au niveau équipement, un plus grand nombre de fonctions peuvent être confiées à des logiciels en remplacement de composants mécaniques ou de logiques pré-câblées.

Ensuite, au niveau sous-système, l'utilisation des divers équipements peut être coordonnée de manière plus fine, en n'utilisant chaque fonction que lorsque nécessaire et en la plaçant en repos le reste du temps. De même, l'intelligence embarquée permet une meilleure gestion des éléments consommables. Enfin, au niveau système, la capacité de prévoir l'évolution de l'environnement et du plan d'opération permet d'harmoniser au mieux la gestion des consommables et de contrôler dans une certaine mesure les conditions d'environnements (thermique essentiellement). L'intelligence embarquée offre en outre la possibilité de reconfigurer automatiquement en vol des équipements défaillants, sans pour autant alourdir sensiblement le satellite.

Néanmoins, ces avantages sont contre-balançés par 2 sortes de limitations :

- la fiabilité intrinsèque des calculateurs de bord, principalement en ce qui concerne les technologies sensibles à l'environnement spatial et les mémoires semi-conducteurs,
- le coût des systèmes informatiques embarqués, qui est essentiellement non récurrent et concerne surtout les logiciels d'interface et de gestion. La définition de ces logiciels devrait être standardisée de manière à mieux répartir l'investissement. Un optimum reste à trouver entre le niveau d'effort de validation au sol, la capacité de reprogrammation en vol et la disponibilité attendue.

12.5 - ENTRETIEN EN ORBITE

L'entretien de satellites en orbite est une solution attractive pour assurer le fonctionnement de satellites ou plates-formes prévues pour de longues ou très longues durées de vie ; cela semble une solution assez souple puisqu'il suffit, en cas de besoin ou systématiquement, d'envoyer un véhicule de dépannage pour remettre à niveau le satellite. En fait, c'est une solution techniquement difficile puisqu'elle fait appel à des techniques nouvelles comme la robotique, le rendez-vous et l'assemblage, le véhicule spécifique d'entretien et qu'elle a des impacts importants sur la conception du satellite lui-même.

12.5.1 - Intérêt de l'entretien en orbite

L'intérêt de l'entretien en orbite des satellites est multiple car, loin de se limiter à la réparation ou à la prévention de pannes, cette solution peut également influencer la stratégie d'utilisation des satellites. Ainsi, outre l'entretien systématique et la réparation, l'utilisation d'un véhicule d'entretien peut permettre de ravitailler le satellite en ergols, de réparer ou échanger tout ou partie de sa charge utile, voire de participer à certains tests de bon fonctionnement. Cette solution écarte ainsi la réduction de plusieurs des causes de limitation de la durée de vie.

12.5.1.1 - Ravitaillement en ergols

La masse d'ergols embarquée au lancement dépend de la durée de la mission et des opérations d'utilisation du satellite. Réciproquement, la durée de vie espérée du satellite en fonction de sa masse d'ergols elle-même limitée par les capacités du lanceur. Ainsi, ces dernières années, la fin de la vie opérationnelle d'un certain nombre de satellites a été due à l'épuisement des ergols (MILLAS, par exemple). Il est donc intéressant d'avoir la possibilité d'amener des ergols au satellite, possible pour prolonger l'utilisation de la plate-forme que pour remédier à un épuisement prématuré en cas d'opérations imprévues consommatrices d'ergols. Dans le cas de satellites de longue durée de vie, il serait également possible d'adapter le concept et la stratégie d'utilisation des ergols à un ravitaillement systématique au bout d'un certain temps (moitié de la durée de vie prévue) afin que le satellite pourrait être dimensionné par une partie (la moitié par exemple) de la durée de la mission, ce qui représente un gain de masse en orbite.

En contre-partie, ravitailler un satellite en orbite implique des opérations complexes et coûteuses des opérations correspondantes : entre autres, le satellite doit être conçu pour permettre un système de propulsion adapté soit à un remplissage de ses réservoirs par un véhicule d'entretien, soit à l'adjonction et à l'utilisation d'un réservoir externe.

Parmi les différents modes de ravitaillement en ergol, on peut distinguer deux modes : transfert des ergols depuis le véhicule d'entretien jusqu'aux réservoirs du satellite ou ajout d'ergol externe. Cette solution suppose un concept adapté du système de propulsion du satellite et de ses équipements et d'équipements (pompe, sur-pressurisation) spécifiques pour un transfert de fluide. Une autre solution consiste à ajouter des réservoirs au satellite (figure 4) ou à remplir les réservoirs vides, ce qui évite le problème de remplissage ; mais l'addition de réservoirs modifie les caractéristiques physiques du satellite et le remplacement de réservoirs conduit à une série d'opérations complexes à effectuer.

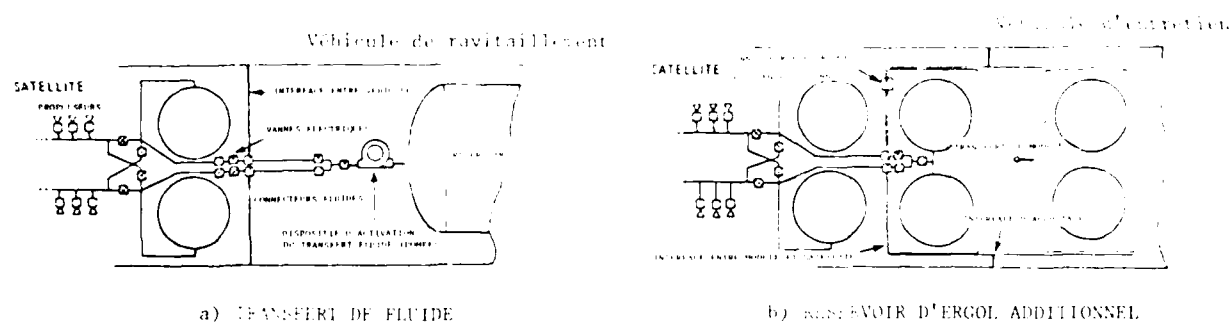


Figure 4 : MODES POSSIBLES DE RAVITAILLEMENT EN ERGOLS

12.5.1.2 - Entretien et réparation

L'entretien systématique du satellite en orbite consiste à échanger les équipements estimés critiques, c'est-à-dire ceux dont la durée de vie est limitative ou inférieure à la durée de vie spécifique du système. Cette possibilité influence considérablement la philosophie de conception du satellite puisqu'on pourrait dissocier la durée de vie des éléments critiques de celle du satellite lui-même. Ces éléments critiques pourraient être dimensionnés non plus pour la durée de vie du système complet, mais pour la période s'écoulant entre le tir du satellite et la visite du véhicule d'entretien ou entre deux visites successives du véhicule d'entretien. Cette période pourrait être optimisée en fonction du choix des technologies des éléments critiques et réciproquement.

Dans ces conditions, la conception du satellite vis-à-vis de sa durée de vie spécifique prendra en compte :

- la fiabilité des équipements à échanger pour la durée d'une période d'entretien, ce qui conduit à des exigences moins sévères sur ces équipements, à des assouplissements de leurs conditions d'utilisation et à un gain de masse,
- la fiabilité des autres équipements sur toute la durée de vie du satellite.

Parmi les équipements qui pourraient bénéficier d'un tel entretien systématique, on peut citer :

- Les batteries : les batteries NiCd en orbite basse ont une durée de vie de quelques années, suivant leurs conditions d'utilisation (qui dépendent du nombre de batteries embarquées). Les batteries NiH₂ semblent avoir une durée de vie plus longue, mais toutes deux pourraient être critiques vis-à-vis d'une durée de vie de 10 à 15 ans en orbite basse. Le remplacement des batteries au bout de quelques années serait une solution attractive et influencerait le choix du type des batteries et ses conditions d'utilisation.
- Les boîtiers électroniques.
- Les paquets de gyroscopes dont la fiabilité se dégrade en fonction du temps, en particulier à cause de l'usure des paliers pour certains types de gyros.
- Les revêtements thermiques passifs qui pourraient être ajoutés dans le cas de dégradation trop importante des caractéristiques thermo-optiques des revêtements multicouches initiaux, ou remplacés (cas des radiateurs) lorsqu'ils sont montés directement sur un équipement lui-même à remplacer.
- Les éléments mécaniques tels pompes, actionneurs...
- Les panneaux solaires dont le remplacement pourrait être envisagé suivant le cas de mission, par exemple si les dégradations subies au cours de la vie sont trop importantes (ce qui dépend des caractéristiques orbitales), ou si l'échange ou l'addition d'une charge utile entraîne un accroissement important de la demande en énergie. Un tel échange aurait essentiellement un impact sur le volume sous coiffe, la masse au lancement et les caractéristiques physiques du satellite en orbite.

Outre l'entretien systématique, la réparation d'équipements défectueux pourrait s'effectuer en orbite, dans la mesure toutfois où la panne peut être détectée et localisée et où l'équipement a été conçu pour être échangé.

Parmi les approches envisageables pour l'entretien en orbite des satellites, l'une consiste à échanger l'équipement en cause, ce qui suppose qu'il ne participe pas au fonctionnement du satellite pendant au moins toute la durée de l'opération.

Une autre approche consiste à ajouter un nouvel équipement, ce qui suppose que le satellite soit capable de prendre en compte un élément additionnel, tant au point de vue alimentation et distribution de données qu'au point de vue caractéristiques physiques.

12.5.1.3 - Modification de la charge utile

Il pourrait être intéressant d'utiliser le véhicule d'entretien pour échanger toute ou partie de la charge utile, soit parce qu'il y a défectuosité, soit pour modifier ou changer la mission tout en conservant la même plate-forme. Cela permet, entre autre, de remédier à l'obsolescence qui est également l'une des causes possibles de limitation de la durée de vie des satellites. Cette modification de la charge utile ne peut s'effectuer que si la plate-forme et la charge utile ont été conçues dans cette éventualité et à condition que la nouvelle charge utile soit compatible avec les ressources offertes par la plate-forme, bien que celle-ci puisse éventuellement être remise à niveau (addition de batteries, par exemple) dans les limites de sa conception.

12.5.1.4 - Impacts de l'entretien en orbite

L'impact de l'entretien en orbite sur la durée de vie et la fiabilité du satellite est représenté figure 5. Elle montre l'accroissement de ces deux paramètres en cas d'entretien seul, de ravitaillement seul et de combinaison entretien plus ravitaillement.

On voit qu'il est avantageux de combiner toutes les interventions, ce en conséquence de planifier les dates d'intervention du véhicule d'entretien au cours de la vie du satellite. Cela permet d'avoir un dimensionnement des réservoirs d'ergols cohérent avec le choix des technologies des éléments critiques, tous deux étant fonction de la date prévue pour le (ou les) entretiens en orbite. De la même façon, il serait intéressant de coupler une modification de la charge utile avec une mission d'entretien.

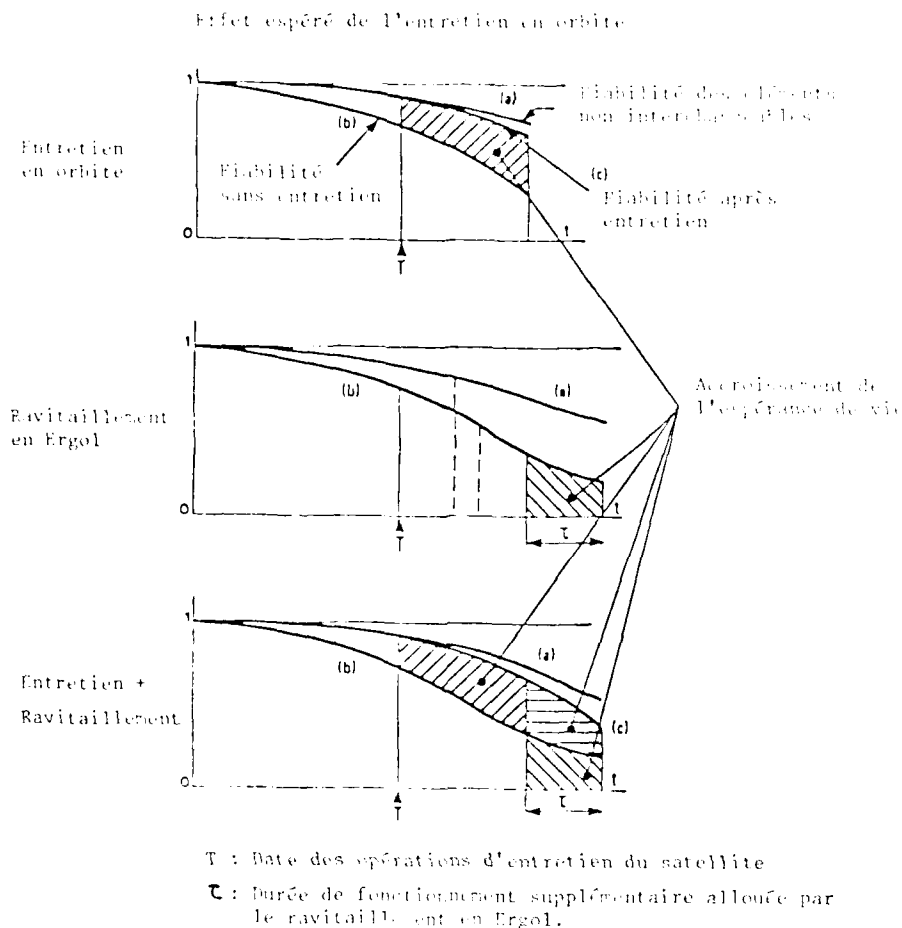


Figure 5 : IMPACT DE L'ENTRETIEN EN ORBITE SUR LA DURÉE DE VIE DU SATELLITE

En contrepartie, l'entretien en orbite impose des règles de conception du satellite. Ainsi :

- Tous les équipements ne pourront pas bénéficier de cet entretien; les équipements concernés doivent être implantés dans la zone d'intervention du véhicule d'entretien (qui diffère suivant que les manipulations sont effectuées par un bras manipulateur ou par l'homme), être d'accès aisé, et avoir des connexions électriques, mécaniques et/ou fluides adéquates avec le reste de la plate-forme.
- Le satellite doit être accessible par un véhicule d'entretien; et donc offrir un interface d'accostage et posséder les équipements de rendez-vous correspondants, les différents sous-systèmes (contrôle d'attitude, thermique, énergie, gestion des données, ...) doivent être compatibles avec les opérations d'entretien.
- Les équipements spécifiques de la mission d'entretien (équipements de rendez-vous et accostage par exemple) doivent être tels que le satellite puisse toujours bénéficier de l'entretien.

On voit donc que, par rapport aux autres solutions proposées pour augmenter la durée de vie des satellites, l'entretien en orbite a des impacts importants sur la philosophie de conception et d'utilisation du satellite.

12.5.2 - Éléments d'infrastructure associés

Il est évident que l'entretien en orbite de satellites ne pourra être assuré sans une infrastructure orbitale associée. Le choix des moyens à mettre en œuvre pour assurer cet entretien en orbite dépend du type d'opérations envisagées et, en particulier, de la présence ou non de l'homme. En effet, les opérations d'entretien en orbite peuvent être effectuées de manière automatique par l'intermédiaire d'un véhicule d'entretien pourvu d'équipements de robotique nécessaires (bras manipulateur, caméras, ...), transportant les éléments de rechange et contrôlé depuis le sol (ou depuis une station spatiale); ces opérations peuvent également bénéficier de l'intervention humaine.

dans le cas d'entretien automatique (figure 6b), le véhicule doit effectuer un rendez-vous avec le satellite (ou la plate-forme) et s'y arrimer ; le contrôle de l'ensemble des deux véhicules et la fourniture des ressources pendant toute la phase de couplage peuvent être effectués par l'un ou l'autre, suivant le cas. Le ravitaillement en ergols suppose la mise en place du réservoir auxiliaire par le bras manipulateur ou la connexion entre le système de propulsion du satellite et l'interface correspondant du véhicule. Les équipements à échanger doivent être dans la zone atteignable par le bras manipulateur. Les opérations de manipulation seront de préférence contrôlées de manière continue par la station sol, ce qui suppose une infrastructure de communication adéquate. Le véhicule peut être lancé par un lanceur conventionnel sur l'orbite du satellite (figure 6b), par un vaisseau type Space Shuttle depuis son orbite (7/c) ou depuis une station spatiale en orbite basse (6/d) ; il peut être récupérable ou consommable.

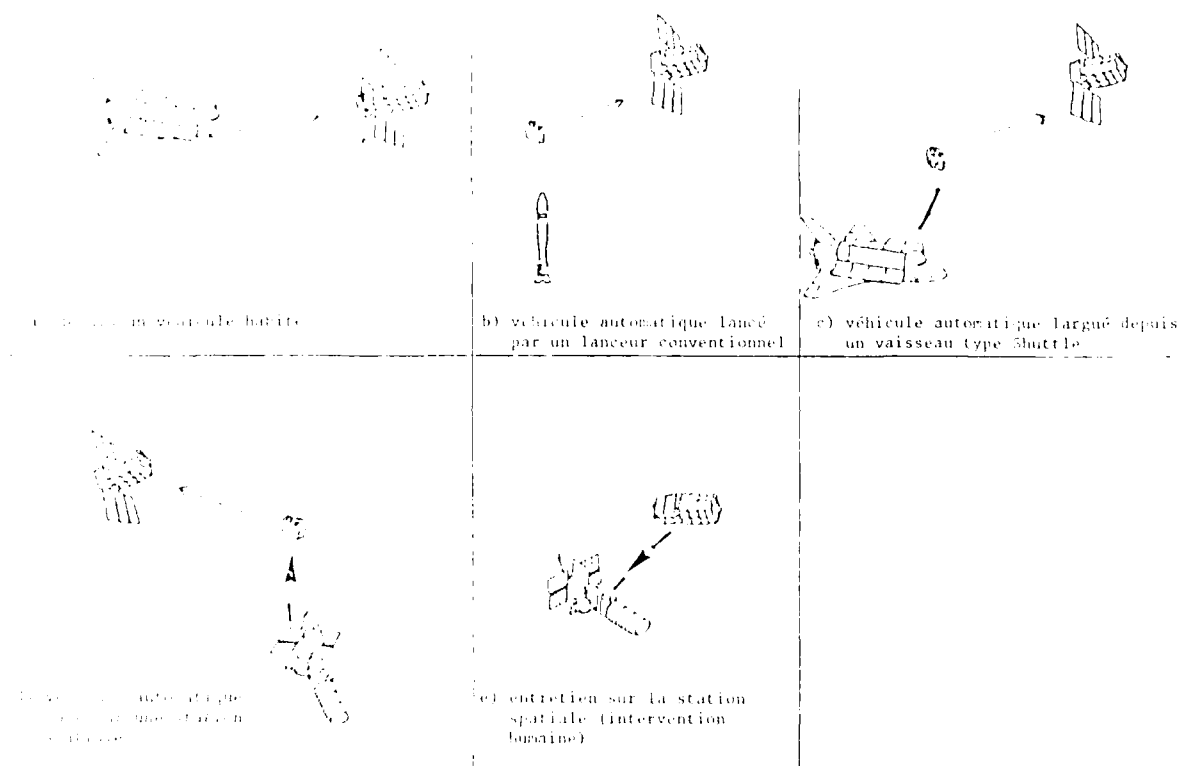


Figure 6 : SCÉNARIOS POSSIBLES D'ENTRETIEN EN ORBITE

Dans le cas d'intervention humaine, les opérations d'entretien pourraient se faire directement à partir d'un véhicule habité, type Space Shuttle (6/d), ou après amarrage à une station spatiale (6/e) en orbite basse. Dans ce dernier cas, un véhicule spécifique (du type véhicule d'entretien) pourrait être utilisé pour aller chercher le satellite en orbite, le ramener à la station spatiale avant d'aller le déposer à nouveau sur son orbite une fois l'entretien terminé. Les opérations d'entretien seraient alors effectuées par l'homme, soit directement, soit par l'intermédiaire d'un manipulateur situé sur la station. L'intervention directe de l'homme permet d'élargir la zone atteignable pour échanger les éléments critiques à tout le véhicule ; mais même avec l'assistance de l'homme, qui permet probablement d'échanger plus d'équipements, il faut encore que le satellite soit conçu pour un tel entretien, et en particulier que les équipements en cause soient accessibles et démontables.

Il reste que le coût de l'entretien en orbite dépendra du type de scénario d'entretien choisi.

12.6 - CONCLUSIONS

La durée de vie des satellites a tendance à augmenter depuis une décennie. Cet accroissement, qui résulte probablement d'une amélioration continue de la technologie, va se heurter aux apports limités par le lanceur, à l'intérêt de la mission au bout d'une longue durée et à la technologie même d'un certain nombre d'équipements qui vont devenir critiques faute de pouvoir remplir les exigences d'une durée de vie accrue.

Plusieurs approches sont envisageables pour remédier à ces problèmes et améliorer ainsi la durée de vie actuelle des satellites. Parmi elles, l'utilisation de l'intelligence de bord permet d'optimiser les conditions de fonctionnement des sous-systèmes et de leurs éléments critiques, en prédisant par exemple leur charge de travail à venir, et d'assurer les reconfigurations nécessaires lors de défaillances d'équipements. Toutefois, cette solution trouve sa propre limitation dans la tenue de ses composants électroniques et dans le nombre fini d'éléments et de ressources embarqués.

La seule possibilité de s'affranchir de cet ensemble fini et de pouvoir intervenir même après lancement consiste à effectuer de l'entretien en orbite. Cette solution consiste à aller remplacer en orbite les équipements critiques et à ravitailler en ergols le satellite. Elle assure donc la prévention (entretien systématique) et la réparation, et ajoute donc un degré de liberté à la conception du satellite car il devient possible de dimensionner la masse d'ergols et de choisir les technologies d'éléments critiques pour des durées de vie plus courtes que celle spécifiée pour le satellite. L'échange de toute ou partie de la charge utile permettrait même de faire évoluer la mission. Toutefois, cette solution a ses contreparties puisqu'elle impose des règles de conception adaptées au satellite et suppose l'existence d'une infrastructure orbitale associée et l'utilisation de nouvelles techniques comme le rendez-vous, l'assemblage, la robotique.

Ces deux solutions sont d'ailleurs complémentaires puisque l'entretien en orbite fera appel à l'intelligence de bord du satellite pour assurer un certain nombre de fonctions. Ces deux techniques auront un rôle important dans l'utilisation d'un satellite à longue ou très longue durée de vie.

PERFORMANCES DES VISEURS D'ETOILE A TUBE DISSECTEUR D'IMAGE : LEURS AVANTAGES ET INCONVENIENTS COMPARES AUX VISEURS D'ETOILE A DETECTEUR SOLIDE

H. BOURCIER, Ingénieur en Chef - M. VITE, Responsable Technique
SODERN - 1, avenue Descartes - 94450 LIMEL BREVANNES, France

Résumé

Les applications spatiales du futur nécessiteront de développer des viseurs d'étoile de hautes performances qui permettront soit de définir avec précision des références d'attitude de plateformes spatiales, soit de mesurer l'attitude du véhicule dans l'espace. Les performances accessibles telles que la magnitude maximale des étoiles détectables, la luminance du fond de ciel admissible, la résolution angulaire, la cadence des mesures dépendent des caractéristiques physiques intrinsèques des détecteurs et de leur mode de fonctionnement.

L'objet de cette publication est, d'une part, de fournir les performances accessibles aux viseurs d'étoile à tube dissecteur d'image en s'appuyant sur les équipements développés et réalisés pour l'IPS/Spacelab et le satellite EXOSAT, d'autre part, d'évaluer leurs avantages et inconvénients comparés aux équipements utilisant des détecteurs solides à transfert de charge. Les performances sont analysées en fonction des propriétés propres aux détecteurs et des contraintes technologiques liées à l'utilisation de chacun d'eux.

PERFORMANCES DES VISEURS D'ETOILE A TUBE DISSECTEUR D'IMAGE

Au cours des années 1977 à 1979, SODERN a développé une génération de viseurs d'étoile de hautes performances pour les besoins de recalibration de l'Instrument Pointing System (IPS) du Spacelab et pour le contrôle d'attitude du satellite scientifique EXOSAT qui a été lancé avec succès le 26 mai 1983.

Afin de tenir les spécifications de précision dans les conditions d'environnement spatial, le concept SODERN comprend l'utilisation d'un tube dissecteur d'image (TDI) avec un système de recalibration optique qui permet d'obtenir un excellent rapport champ/précision.

Principe de fonctionnement du viseur d'étoile pour IPS

Une optique focalise le flux provenant de l'étoile à poursuivre sur la photocathode du tube (fig. 1).

Les électrons venant de la photocathode sont accélérés, focalisés et déviés par la grille d'accélération et un ensemble de trois bobines ce qui permet de conjuguer n'importe quel point de la photocathode avec un trou placé en face de la première dynode d'un multiplicateur d'électrons. Ce trou définit la dimension du champ de vue instantané du viseur (CVI).

Le taux de comptage des photons venant de l'étoile sur l'endroit choisi de la photocathode est mesuré par un compteur connecté à la sortie d'un comparateur.

Pour l'application IPS, le taux de comptage l'obscurité est inférieur à 20 coups/seconde tandis que le taux de comptage pour l'étoile la moins brillante est supérieur à 2500 coups/seconde.

Mode recherche

A partir d'une position envoyée par télécommande, qui peut se trouver n'importe où dans le champ, le capteur balaye le champ de vue en suivant un motif de recherche présenté en figure 2.

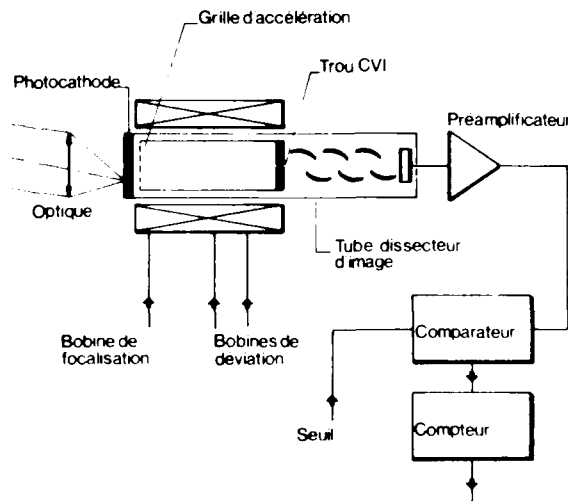


Figure 1. Principe du viseur d'étoile IPS

Pour chaque pas élémentaire, le taux de comptage N_i de la position i pour un temps d'intégration T est comparé au taux de comptage $N(i-2)$ de la position $i-2$. Si la différence dépasse un seuil qui est fonction de la magnitude de l'étoile recherchée, une nouvelle mesure de confirmation $N_{ic} - N(i-2)c$ est effectuée ; si cette différence dépasse le seuil, l'étoile est détectée et le senseur passe automatiquement en mode poursuite.

Pour une probabilité de détection (P_D) et une probabilité de fausse détection (P_{FD}) déterminées, le temps élémentaire (T_E) d'une position du mode recherche est donné par :

$$T_E = \frac{P^2}{N_e} \left(1 + 2 \frac{N_f}{N_e} \right)$$

où :

N_e est le taux de comptage étoile en coups/s

N_f est le taux de comptage lié à la luminance du fond de ciel en coups/s

P est un coefficient qui dépend des probabilités P_{FD} et P_D .

La durée totale d'un balayage de recherche correspond au produit du nombre total de pas N par le temps élémentaire T_E ; N est approximativement proportionnel au carré du rapport champ de recherche ϕ_R sur champ de vue instantané ϕ_{CVI} .

Mode poursuite

Lorsqu'une étoile est détectée, un motif de balayage à quatre positions (fig. 3a) est utilisé. Un système en boucle fermée permet d'asservir le motif de balayage de sorte que :

$$N_1 - N_2 = N_3 - N_4 = 0,$$

ce qui fait coïncider le centre du motif de balayage avec le centre énergétique de l'image de l'étoile. La mesure des courants qui traversent les bobines de déviation du tube dissecteur permet de connaître la position de l'image de l'étoile sur la photocathode (coordonnées bobines).

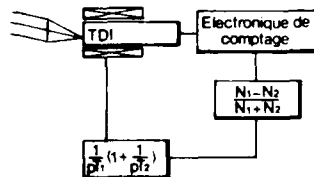


Figure 3a. Principe du mode poursuite.

La figure 3b représente la réponse de la boucle de poursuite suivant l'axe de balayage : N_e est le taux de comptage étoile en c/s (2500 c/s pour $\lambda = 0,45 \mu$ et $W = 1$), la distance angulaire correspondant à la partie linéaire de la courbe de réponse.

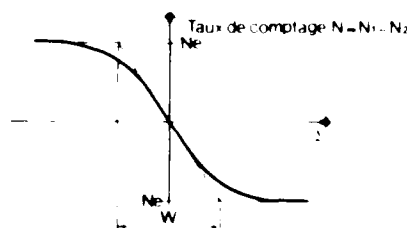
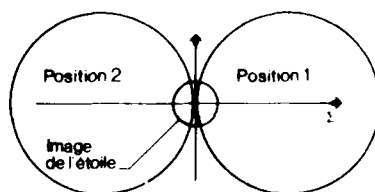


Figure 3b. Réponse de la boucle de poursuite.

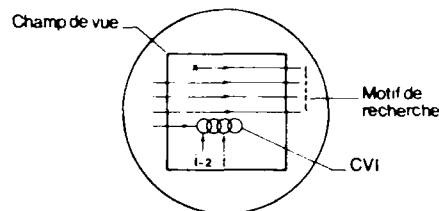
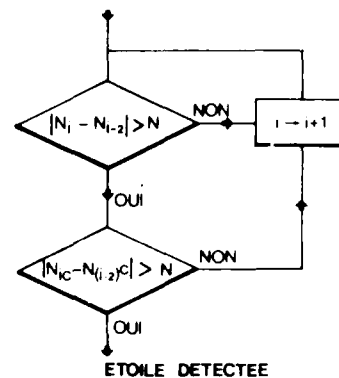
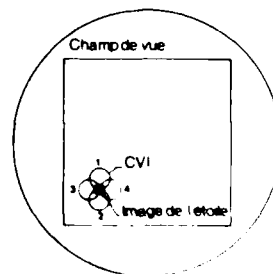


Figure 2. Principe du mode recherche.



Si T_M est la durée de mesure l'objet de balayage et N_p le taux de comptage dû au fond de ciel, la résolution angulaire équivalente au bruit, RA, est donnée par la relation :

$$(RA)^2 = \frac{W^2}{N_p T_M} \cdot (1 + \frac{N_f}{N_p})$$

Système de recalibration optique

La précision pour ce type de configuration dépend :

- de l'effet du champ magnétique sur l'électronique du tube dissecteur d'image,
- des effets chromatiques de l'optique,
- de la stabilité du montage mécanique optique-TBI par rapport aux points de fixation du viseur,
- de la stabilité de l'électronique, principalement, pour le courant de focalisation, la tension d'accélération de la grille et la conversion digitale-analogique des convertisseurs fournissant les courants aux bobines de déviation,
- de la non linéarité et de la stabilité de la loi de déviation en fonction des courants qui circulent dans les bobines.

Les effets chromatiques peuvent être corrigés par la définition de l'optique et les effets du champ magnétique alternatif par l'introduction d'un blindage approprié.

Pour les autres effets, il est difficile d'atteindre une linéarité et une stabilité compatibles avec un objectif de 2.10^4 pour le rapport champ/précision.

C'est pourquoi le viseur d'étoile est équipé d'un système de recalibration optique qui projette un ensemble de 400 (20×20) étoiles de référence dont les directions sont parfaitement connues et stables par rapport à l'axe optique du viseur qui est défini par des miroirs d'alignement intégrés à l'optique du viseur (fig. 4).

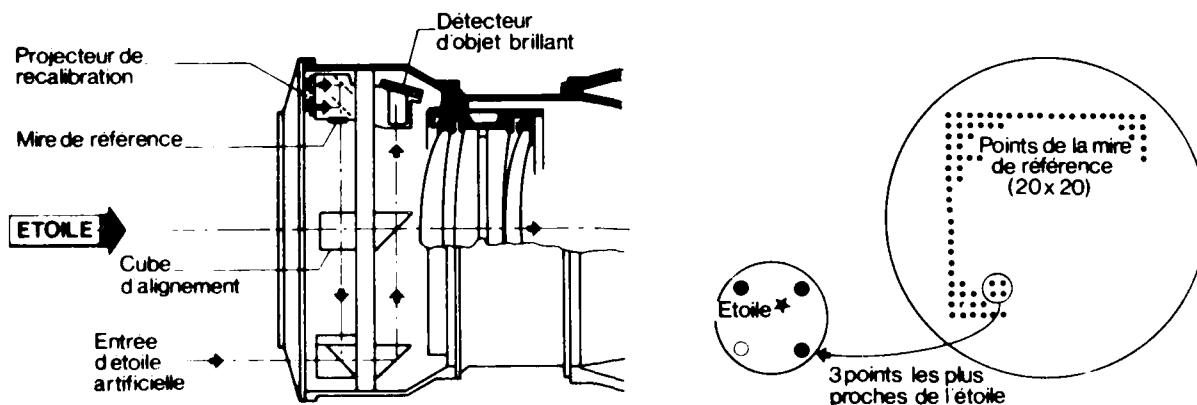


Figure 4. Principe du système de recalibration optique

Le flux lumineux provenant de ces étoiles de référence est modulé afin d'éviter les interférences mutuelles avec les étoiles réelles. Le viseur mesure alternativement les coordonnées bobines de l'étoile réelle poursuivie et les coordonnées bobines des trois étoiles artificielles qui sont les plus proches de l'étoile réelle. Les mesures sur les étoiles artificielles permettent de calculer la loi de transfert de passage des coordonnées bobines aux coordonnées géométriques qui sont directement corrélées aux miroirs de référence du viseur. L'application de cette loi de transfert aux coordonnées bobines de l'étoile permet de délivrer les coordonnées géométriques de l'étoile réelle. La durée de mesure nécessaire au calcul de la loi de transfert des coordonnées bobines aux coordonnées géométriques est de 36 secondes.

Le processus de correction se déroule continuellement de sorte que les dérives qui ont des fréquences d'évolution inférieures seront automatiquement corrigées par l'utilisation du système de recalibration optique.

L'efficacité de ce système de recalibration optique dépend du nombre de points du projecteur d'étoiles artificielles.

Pour un tube ITT F 4012 RP utilisé sur un diamètre de photocathode de 14 mm, la figure 5 montre l'écart, en microns, entre la position réelle de l'étoile et la position mesurée à 1 écart-type en fonction du nombre d'étoiles artificielles du système de recalibration optique.

Pour le viseur d'étoile SED 04 développé pour l'IPS du SPACELAB qui utilise un ensemble de 20 x 20 points de mire de référence, le rapport champ/précision à 3 écarts-type ressort à $2,5 \cdot 10^4$, ce qui est compatible avec l'objectif spécifié de : $2''/0,75 \text{ arc-sec} = 9600 (3\sigma)$ qui comprend en outre les effets des dérives thermiques, des distorsions chromatiques et des variations de champ magnétique.

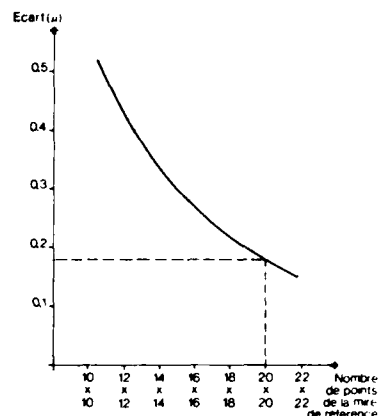


Figure 5. Précision de recalibration en fonction du nombre de points de mire

Principales caractéristiques et résultats expérimentaux du viseur d'étoile SED 04 développé pour l'Instrument Pointing System du Spacelab

Les principales caractéristiques du viseur SED 04 sont résumées dans le tableau 1.

Caractéristiques SED 04	
champ de vue carré	2° x 2°
optique longueur focale diamètre d'ouverture	265 mm 80 mm
détecteur tube dissecteur d'image photocathode champ de vue instantané	ITT 4012 RP S 20 200μm/160 sec
électrique cadence des informations fréquence de fonctionnement du microprocesseur mémoire PROM mémoire RAM puissance (20 à 32 V)	6 x 17 bit/sec 16 MHz 14 K bytes 2 K bytes 21 W
dimensions et masse : tête optique - longueur - diamètre - masse	550 mm 210 mm 7,5 kg
boîtier électronique - longueur - largeur - hauteur - masse	245 mm 210 mm 160 mm 4,5 kg
* Mode étoile . magnitude visuelle Mv . mode dégradé . température de couleur . luminance du fond de ciel	+ 2 à + 8 + 9 3000 à 28000K 8,4 μlm/m ² sr
* Mode 1 étoile . temps d'acquisition maximal . poursuite : - bande passante - angle équivalent au bruit	76 s 2,2 Hz 0,5 sec/√Hz
* Mode 2 étoiles . temps d'acquisition . poursuite : - bande passante - angle équivalent au bruit	152 s 1,1 Hz 0,75 sec/√Hz
Linéarité et stabilité (à 3 écarts types)	0,75 sec(3σ)

Tableau 1 - Principales caractéristiques du viseur d'étoile SED 04

Les principaux résultats expérimentaux sont résumés dans le tableau 2. Ces résultats sont représentatifs des mesures effectuées sur les 6 viseurs d'étoiles SED 04 produits entre 1979 et 1983.

Caractéristiques SED 04	
Mode 1 étoile taux de comptage pour magnitude visuelle + 8 taux de comptage pour un fond de ciel de 8,4 $\mu\text{lm}/\text{m}^2 \text{ sr}$. temps élémentaire d'acquisition . temps d'acquisition maximal . angle équivalent au bruit pour une étoile $M_v = 8$ et fond de ciel	2800 c/s 500 c/s $6,10^{-3} \text{ s}$ 70 s $0,45 \text{ sec} / \sqrt{\text{Hz}}$
Mode 2 étoiles . magnitude visuelle + 8 . temps d'acquisition maximal . angle équivalent au bruit pour une étoile $M_v = 8$ et fond de ciel	2800 c/s 132 s $0,7 \text{ sec} / \sqrt{\text{Hz}}$
Stabilité . linéarité de réponse en fonction de la position dans le champ de vue, pour 28°C et une température couleur . écart dû aux différentes températures de couleur dans la gamme (3000-28000 K) . écart dû aux variations de température dans la gamme (28 - 32°C) pour la tête optique, et dans la gamme (+ 10 + 50°C) pour le boîtier électronique . influence champ magnétique 0,4 gauss < 3 Hz Stabilité totale (somme quadratique)	0,45 sec 3σ 0,2 sec 3σ 0,1 sec 3σ 0,1 sec 3σ 0,52 sec 3σ
Rapport champ/précision sur chaque axe (à 3 écarts types)	13850

Tableau 2 - Résultats expérimentaux du viseur d'étoile SED 04

EVALUATION DES AVANTAGES ET INCONVENIENTS COMPARES DES VISEURS D'ETOILE A TUBE DISSECTEUR D'IMAGE ET A DETECTEUR SOLIDE DE TYPE CCD

Les principaux paramètres qui caractérisent les performances d'un viseur d'étoile sont :

- la résolution angulaire équivalente au bruit pour l'étoile la moins brillante en fonction de la cadence des mesures,
- la durée d'acquisition pour l'étoile la moins brillante,
- le rapport champ/précision.

L'évaluation comparative pour ces paramètres est faite à partir du tube dissecteur d'image F 4012 RP de ITT ayant une photocathode S.20 et d'un détecteur solide SID 52501 de RCA.

Résolution angulaire équivalente au bruit

Pour un dissecteur d'image, la résolution angulaire équivalente au bruit (RA) est donnée par :

$$(RA)_{\text{INT}}^2 = W^2 \cdot \frac{1}{\left(\frac{S}{R}\right)^2}$$

où W est la distance angulaire correspondant à la partie linéaire de la boucle de poursuite en seconde d'arc,

et $\left(\frac{S}{B}\right)^2 = \frac{N_e T}{1 + 2 \frac{N_f}{N_e}}$ où T est la durée unitaire d'une boucle de poursuite en seconde
 N_e le taux de comptage étoile,
 N_f le taux de comptage du fond de ciel

Pour un champ de mesure circulaire C, exprimé en degrés, un nombre d'ouverture optique N, un diamètre utile de photocathode de 14 mm, une bande spectrale de mesure correspondant à une photocathode S.20, une distance angulaire W de 40µm, un champ de vue instantané de 75µm, une luminance de fond de ciel L_f en candela/m² et une étoile dont la température de couleur est de 6000 K, la résolution angulaire équivalente au bruit en seconde d'arc est donnée par la relation :

$$(RA)_{IDT}^2 = \frac{1}{T} \cdot 1,9 \cdot 10^{-7} C^4 N^2 2,51 mV \left[1 + 5,5 \cdot 10^{-3} C^2 L_f 2,51 mV \right]$$

Pour un détecteur solide de type CCD exploité en mode poursuite sur une zone élémentaire carrée de 3 x 3 pixels et utilisant une tache image circulaire ayant un diamètre de deux pixels, la résolution angulaire équivalente au bruit (RA) est donnée par :

$$(RA)_{CCD}^2 = (\text{dimension angulaire d'un pixel})^2 \cdot \frac{1}{\left(\frac{S}{B}\right)^2}$$

où la dimension du pixel est exprimée en seconde d'arc.

En négligeant le bruit de lecture du CCD, le rapport signal/bruit est donné par :

$$\left(\frac{S}{B}\right)^2 = \frac{N_e T}{1 + 9 \frac{N_{obs}}{N_e} + 9 \frac{N_f}{N_e}}$$

où T est la durée d'intégration élémentaire prise identique à la durée d'une boucle élémentaire de poursuite pour le tube dis-secteur

N_e est le signal étoile en électrons/s

N_{obs} est le courant d'obscurité en électrons/s

N_f est le signal dû au fond de ciel sur un pixel en électrons/s.

Pour un champ de mesure carré, de diagonale C_0 exprimée en degrés, un nombre d'ouverture optique N_0 , un pixel carré de 30µm de côté, une bande spectrale de mesure de 0,65µm à 0,9µm, une étoile dont la température de couleur est de 6000 K, une luminance de fond de ciel L_f en candela/m², un courant d'obscurité $N_{obs} = 12000 \cdot 2^{\theta/7}$, θ étant la température de fonctionnement du CCD, la résolution angulaire équivalente au bruit est donnée par la relation :

$$(RA)_{CCD}^2 = \frac{1}{T} \cdot 0,29 \cdot 10^{-7} C_0^4 N_0^2 2,51 mV \left[1 + 6,5 \cdot 10^{-3} C_0^2 L_f 2,51 mV + 4 \cdot 10^{-5} \cdot 2^{\theta/7} \cdot C_0^2 N_0^2 2,51 mV \right]$$

Résolution angulaire comparée pour un fond de ciel nul : $L_f = 0$

$$\text{Le rapport } R = \frac{(RA)_{CCD}^2}{(RA)_{IDT}^2} \text{ est donné par :}$$

$$R = 0,15 \left(\frac{N_0}{N}\right)^2 \cdot \left(\frac{C_0}{C}\right)^4 \left[1 + 4 \cdot 10^{-5} \cdot 2^{\theta/7} \cdot C_0^2 N_0^2 2,51 mV \right]$$

Dans la gamme de magnitude visuelle allant de $M_v = 1$ à $M_v = 8$, on peut approcher la relation champ-magnitude par la relation :

$$M_v = 10,2 - 4,1 \log C \quad \text{où } C \text{ est le champ circulaire en degrés}$$

qui assure la présence d'au moins une étoile dans le champ de vue C quelle que soit l'orientation dans la voûte céleste.

La comparaison pour les 2 détecteurs doit se faire pour les angles solides équivalents afin de respecter la relation champ/magnitude. Le champ carré $C_0/\sqrt{2} \times C_0/\sqrt{2}$ du détecteur solide doit être équivalent au champ circulaire C du détecteur d'image ; les 2 champs C et C_0 sont donc liés par la relation :

$$\left(\frac{C_0}{C}\right)^2 = \frac{\pi}{2}$$

Afin de rester dans un domaine de définition d'optique réaliste pour des champs de vue allant de 2° à 12° , on fait l'hypothèse que le champ et l'ouverture sont liés par la relation :

$$C.N = \text{constante} = 14$$

Le rapport $\frac{N_0}{N}$ est pris égal à 1,4 afin de tenir compte du fait que dans son principe le tube dissecteur n'est que peu affecté par une déformation de la tache image dans le champ. Dans ces conditions, le rapport R devient :

$$R = 0,15 \cdot (1,4)^2 \cdot \frac{\pi}{4} [1 + 4 \cdot 10^{-5} \cdot 2^{0/7} \cdot (14)^2 \cdot 2,51 \text{mv}]$$

En fonction du champ de mesure, R devient :

$$R = 0,74 [1 + 0,78 \cdot 10^{-2} \cdot 2^{0/7} \cdot 2,51^{10 \cdot 2} - 4,1 \log c]$$

Le réseau de courbes présenté en figure 6 explicite l'évolution du rapport R en fonction du champ de vue et de la température de fonctionnement du CCD. Il montre que pour obtenir des résolutions angulaires comparables pour les deux types de détecteurs, il est impératif de refroidir le CCD à une température inférieure à -20°C et que cette contrainte de refroidissement est d'autant moins sévère que le champ de vue du viseur augmente. Néanmoins, dès que la température de fonctionnement du CCD dépasse -20°C les résolutions angulaires pour les deux types de détecteurs ne sont plus comparables car le rapport R croît très vite avec la température.

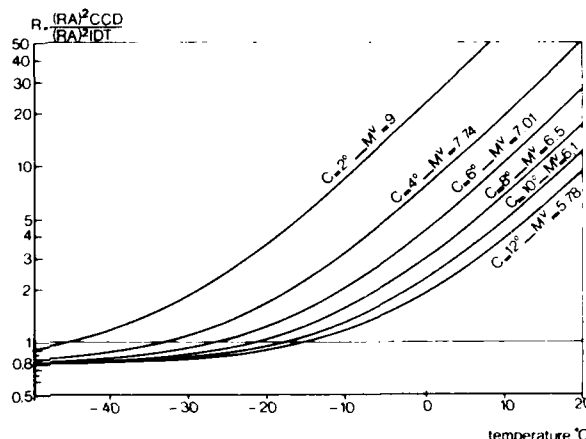


Figure 6. Résolution angulaire comparée des viseurs à CCD et TDI en fonction du champ de vue et de la température du CCD

Résolution angulaire comparée pour un fond de ciel fort

Afin de découpler les paramètres, on suppose ici que le CCD est refroidi à une température de -40°C afin de négliger l'influence du courant d'obscurité.

Le rapport R s'écrit :

$$R = 0,74 \cdot \frac{1 + 6,5 \cdot 10^{-3} C_0^2 L_f 2,51 \text{mv}}{1 + 5,5 \cdot 10^{-3} C^2 L_f 2,51 \text{mv}} \approx 0,74 \cdot \frac{6,5}{5,5} \cdot \left(\frac{C_0}{C}\right)^2 = 1,4$$

On peut en conclure que pour des luminances de fond de ciel fortes, les deux types de détecteurs sont équivalents.

Durée d'acquisition

Pour les deux types de détecteurs, la première phase d'acquisition comprend la détection des positions dans le champ de vue pour lesquelles le signal dépasse un certain seuil qui dépend de la magnitude attendue de l'étoile. Dans cette phase de détection, pour assurer que les probabilités de fausse détection P_{FD} (seuil/bruit supérieur à une certaine valeur) et de non détection P_{ND} (signal-seuil/bruit supérieur à une certaine valeur) sont compatibles avec les besoins de mission, il est nécessaire de remplir la condition :

$$\frac{\text{signal}}{\text{bruit}} \geq P, P \text{ dépendant de } P_{FD} \text{ et } P_{ND}$$

Dans la mesure où le CCD est refroidi à une température de -40°C qui permet de négliger la contribution du courant d'obscurité sur le rapport signal/bruit, la durée d'acquisition pour le CCD sera très inférieure à celle nécessaire pour le tube dissecteur.

En effet, elle correspond approximativement au temps d'acquisition élémentaire pour le tube dissecteur d'image car le rapport signal/bruit est la même pour les deux détecteurs.

Si ϕ_c est le diamètre du champ de vue total et ϕ_{cvi} le diamètre du champ de vue instantané pour le tube dissecteur :

$$T_{\text{acquisitionIDT}} = \frac{\phi_c^2}{2 \phi_{cvi}} \cdot \text{Telementaire}$$

$$T_{\text{acquisitionCCD}} = \text{Telementaire}$$

Il faut cependant remarquer que pour la majorité des applications de hautes précisions le domaine de recherche peut être limité car l'orientation grossière dans la voûte céleste est connue et peut donc être utilisée afin de limiter au strict nécessaire le nombre de pas de recherche pour un viseur à dissecteur d'image.

Rapport champ-précision

Les résultats théoriques et expérimentaux montrent qu'en utilisant un projecteur d'étoiles artificielles de 400 points sur la photocathode d'un tube dissecteur d'image, le rapport champ/précision qui en résulte est d'environ $2,5 \cdot 10^4$. L'utilisation du système de recalibration optique permet de s'affranchir des variations de la répartition d'énergie dans la tache image en fonction de la position dans le champ et des variations de la dimension de la tache image produite par une défocalisation liée au déplacement axial relatif photocathode-optique. En effet, ces fluctuations sont prises en compte par la mesure des étoiles artificielles de mire car dans un domaine de champ délimité par un carré de 4 points de mire successifs, les variations résiduelles sont négligeables. Seule la résolution angulaire est un peu affectée car elle dépend de la pente de la fonction de transfert de la boucle de poursuite qui dépend linéairement de la dimension de la tache image.

Le rapport champ/précision de $2,5 \cdot 10^4$ n'est pas une limite, en effet il peut être amélioré soit en augmentant le nombre d'étoiles artificielles du système de recalibration optique, soit en utilisant un algorithme de correction supplémentaire prenant en compte les distorsions géométriques locales liées aux non linéarités des bobines de déviation.

Un objectif de rapport champ/précision de $5 \cdot 10^4$ à $8 \cdot 10^4$ est tout à fait envisageable sur un équipement opérationnel.

Pour un détecteur solide, il est nécessaire d'utiliser un processus d'interpolation qui conduit à distribuer le flux du signal sur plusieurs pixels et à utiliser les réponses fournies par les pixels individuels afin de déterminer la position géométrique de la tache image en fraction de pixel. Les limitations apportées par la non uniformité de réponse d'un pixel, les écarts de position et de dimension géométrique des pixels, les fluctuations de la répartition d'énergie dans la tache image en fonction de la position dans le champ et de la composition spectrale du rayonnement conduisent à une limitation de la qualité de l'interpolation de l'ordre de 1/30ème à 1/100ème de pixel en fonction de la complexité du traitement et de la calibration (modélisation de la réponse d'un pixel, mémorisation de la sensibilité moyenne pour tous les pixels, corrections des défauts géométriques par utilisation d'une régression linéaire...). Ceci conduit à un rapport champ/précision dans la gamme 7500-30000 sur le plan théorique avec les CCD actuels.

La tenue de ces performances nécessite d'une part le refroidissement du détecteur solide et, d'autre part, soit d'utiliser des corrections prenant en compte les paramètres physiques du détecteur (courant d'obscurité et sensibilité de tous les pixels), soit d'utiliser un composant pour lequel les écarts entre ces paramètres physiques sont inférieurs à 1 % sur l'ensemble du détecteur. La stabilité du positionnement relatif du détecteur solide et de son optique associée est elle aussi très critique. En effet, un déplacement relatif transverse de $\pm 0,5 \mu\text{m}$ correspond à un défaut d'alignement équivalent de 1/30ème de pixel qui ne peut être corrigé que par l'utilisation d'un système de recalibration optique ayant au moins trois étoiles artificielles. Les variations du positionnement axial introduisent des variations de la répartition d'énergie dans la tache image qui sont d'autant plus critiques que l'optique utilisée est très ouverte.

Toutes ces contraintes de répartition d'énergie dans la tache image, de conception thermo-mécanique et d'évolution des paramètres de calibration (courant d'obscurité, sensibilité de tous les pixels) en vieillissement et sous contraintes de rayonnement en ambiance spatiale sont difficilement compatibles avec les besoins d'un équipement opérationnel ; ceci conduit à penser que pour un équipement opérationnel de structure simple, une capacité d'interpolation meilleure que le 1/10ème de pixel, donc un rapport champ/précision de l'ordre de 2000 à 4000, est une limite raisonnable.

Pour améliorer le paramètre champ/précision, il est nécessaire d'une part d'utiliser un système de recalibration optique, d'autre part de développer des méthodes de défocalisation de la tache image qui permettent de rester insensibles aux variations dans le champ de vue et aux variations de répartition d'énergie dans la tache pour les déplacements relatifs axiaux optique-détecteur. Une solution qui semble bien adaptée consiste à utiliser une défocalisation électronique introduite par un tube intensifi-

cteur d'image qui permet, par ailleurs, d'utiliser une optique faiblement ouverte afin de résoudre les problèmes de répartitions d'énergie dans la tache image qui sont purement liés à l'optique. La contrainte d'homogénéité de sensibilité relative des pixels (meilleure que 1 %) reste un paramètre prépondérant de la qualité d'interpolation; une solution consistant à utiliser un détecteur solide illuminé par l'arrière doit permettre d'améliorer la qualité de l'homogénéité de la réponse sur l'ensemble du détecteur.

Ce type de concept, lié à l'utilisation d'une optique faiblement ouverte, introduit par définition une diminution du rapport signal/bruit et nécessite de développer des photocathodes ayant un rendement quantique et une sensibilité améliorés dans la bande 600-900 nm (par exemple, des photocathodes AsGa) ; ce type de photocathode, une fois développée, pourrait elle aussi être intégrée dans un tube dissecteur d'image.

Une approche de solution pour un viseur d'étoile de hautes performances à détecteur solide CCD conduit à la structure suivante :

- détecteur solide, ayant une homogénéité de réponse meilleure que 1 %, illuminé par l'arrière et thermostaté vers 0°C (l'apport du gain du tube intensificateur permet de limiter l'effet du courant d'obscurité par la diminution du temps d'intégration),
- couplage détecteur solide-tube intensificateur,
- optique faiblement ouverte comprenant un système de recalibration optique à trois étoiles artificielles,

dont les performances objectives en environnement spatial seraient aux environs de 1/50ème de pixel, donc un rapport champ/précision de 10000 à 20000.

IMAGE DISSECTOR TUBE STAR TRACKERS PERFORMANCE : ADVANTAGES AND DRAWBACKS AS COMPARED TO THOSE OF CHARGE TRANSFER DEVICE STAR TRACKERS

H. BOURCIER, Ingénieur en Chef - M. VITE, Responsable Technique
SODERN - 1, avenue Descartes - 94450 LIMEIL BREVANNE, France

Abstract

Future space applications will require developing high performance star trackers to provide with either high accurate attitude reference or measurement for spaceborne platforms. The achievable characteristics such as detectable star maximum magnitude, allowable sky background radiance, angular resolution, time resolution depend on detectors intrinsic physical characteristics and internal operation.

The aim of this paper is, on one hand, to give performance characteristics which can be achieved with image dissector tube star trackers based on equipment manufactured for IPS/Spacelab and Exosat satellite, on the other hand, to make an assessment of their advantages and drawbacks with respect to equipment based on solid state devices of charge transfer type. The analysis of the performances is presented versus the intrinsic characteristics of the detectors and the technological constraints related to their use.

PERFORMANCES OF IMAGE DISSECTOR TUBE STAR TRACKERS

During the years 1977 to 1979, SODERN has developed a high performance star trackers generation for the recalibration needs of the Instrument Pointing System (IPS) of Spacelab and for the attitude control of the EXOSAT scientific satellite which has been launched with success on May 26, 1983.

In order to hold the accuracy specifications in space environment conditions, the SODERN concept uses an image dissector tube (IDT) together with an optical recalibration system which allows to achieve an excellent field/accuracy ratio.

IPS star tracker working principle

An optical device focuses light coming from the star to be tracked on the tube photocathode (fig. 1).

Electrons coming from the photocathode are accelerated, focused and deflected by an acceleration mesh and a set of three coils, in such a way that any point of the photocathode may go through a hole facing the first dynode of an electron multiplier. This hole determines the tracker Instantaneous Field of View (IFOV).

The rate of photons coming from the star on a selected part of the photocathode is measured by a counter connected to the comparator output.

Concerning the IPS application, the dark count rate is lower than 20 counts per second and the counting rate corresponding to the dimmest star to be tracked is 2500 counts per second.

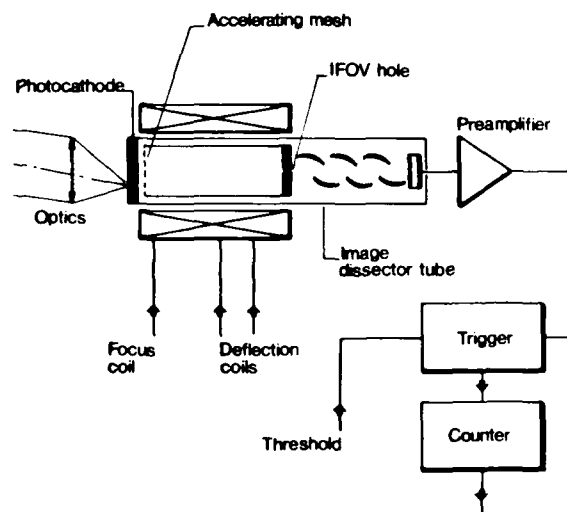


Figure 1. IPS star tracker principle

Search mode

Starting from a point which position can be set by commands anywhere in the field of view, the sensor scans the field of view along to a search pattern shown on figure 2.

AD-P003 382

For each elementary step, the N_1 counting rate of the 1 position for a T integration time is compared to the counting rate N (1-2) of the 1-2 position. If the difference exceeds a given threshold in relation with the searched star magnitude, a new confirmation measure $N_{1c} - N_{(1-2)c}$ is undergone; if this difference exceeds the threshold, the star is detected and the sensor automatically starts working in a tracking mode.

For a (P_D) probability of detection and a (P_{FD}) probability of false detection set in advance, the (T_E) elementary time of a search mode position is given by :

$$T_E = \frac{p^2}{N_e} \left(1 + 2 \frac{N_f}{N_e} \right)$$

where :

N_e is the star counting rate in counts/s

N_f is the counting rate related to sky background radiance in counts/s

p is a coefficient depending on P_{FD} and P_D probabilities.

The total duration of a search scanning corresponds to the product of the total number of N steps by the T_E elementary time; N is approximatively proportional to the square of the ϕ_R search time on ϕ_{IFOV} instantaneous field of view ratio.

Tracking mode

When a star is detected, a four-position scanning pattern is used (fig. 3a). A closed loop system allows to drive the scanning pattern in such a way that :

$$N_1 - N_2 = N_3 - N_4 = 0,$$

that makes the scanning pattern centre coincide with the image or star energy centre. Measurement of currents going through the dissector tube deflection coils allows to know the star image position on the photocathode (coils coordinates).

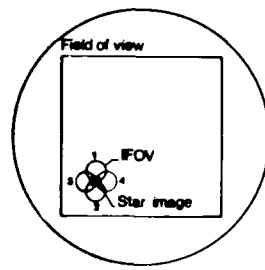


Figure 3a. Tracking mode principle

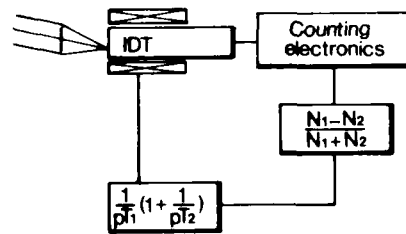


Figure 3b shows the tracking loop response following a scanning axis : N_e is the star counting rate in c/s (2500 c/s for $M_v = 8$) and W is the angular distance corresponding to the linear part of the response curve.

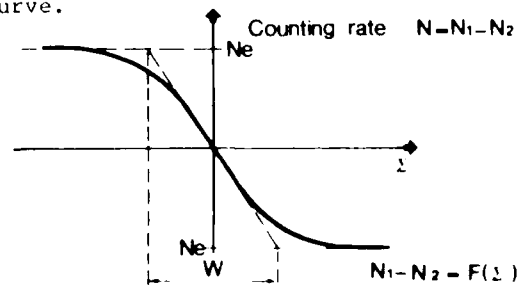
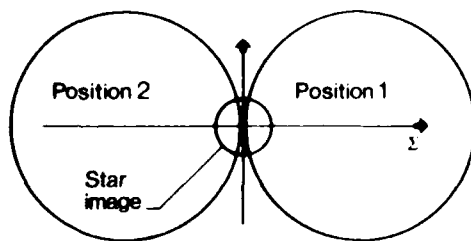


Figure 3b. Tracking loop response

If T_M is the measure duration of a scanning pattern and N_F the counting rate due to sky background, the angular resolution equivalent to noise RA is given by :

$$(RA)^2 = \frac{\omega^2}{N_e T_M} \cdot \left(1 + 2 \frac{N_F}{N_e}\right)$$

Optics recalibration system

Accuracy for this type of configuration depends on :

- magnetic field effect on image tube dissector electronical optics,
- optics chromatic effects,
- stability of IDT-optics mechanical mounting with regard to tracker fixation points,
- electronics stability, especially for focusing current, mesh acceleration voltage and digital to analog conversion of converters providing current to deviation coils,
- non-linearity and stability of deviation law in terms of currents circulating in the coils.

Chromatic effects may be corrected by optics definition and effects of alternative magnetic field by introducing an appropriate shielding.

For other effects, it is difficult to reach a linearity and a stability matching a 2.10^4 objective for the accuracy/field ratio.

That is why the star tracker is equipped with a recalibration optical system which projects a set of 400 (20×20) reference stars which directions are perfectly known and steady with respect to the tracker optical axis which is defined by alignment mirrors integrated in the tracker optics (fig. 4).

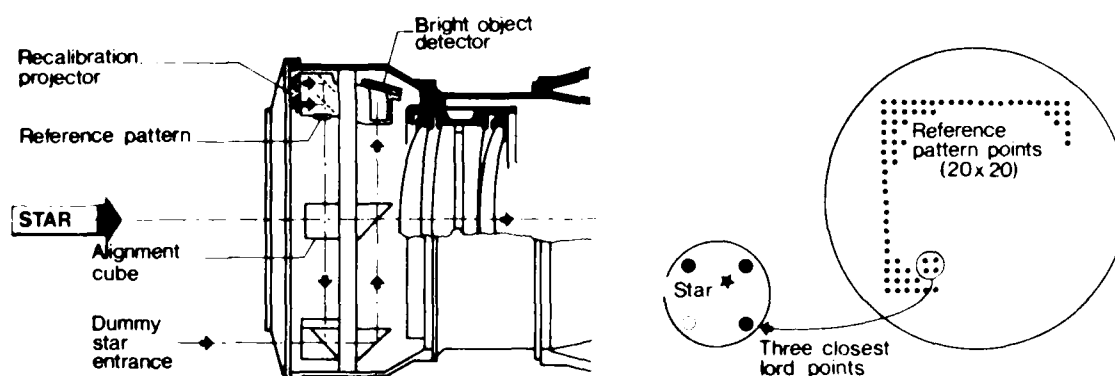


Figure 4. Optical recalibration system principle

Light flow coming from these reference stars is modulated in order to avoid mutual interferences with real stars. The tracker alternatively measures coils coordinates of the real star tracked and the coils coordinates of three artificial stars which are the nearest to the real star. Measures on artificial stars lead to calculate the transfer law of coils coordinates to geometrical coordinates which are directly related to tracker reference mirrors. The application of this transfer law to coils coordinates of the star allows to deliver geometrical coordinates of the real star. The measure time necessary to calculate the transfer law of coils coordinates to geometrical coordinates is 36 seconds.

This correction process works continuously on fast units, which are lower evolution frequencies, are automatically corrected by using an optical recalibration system.

The efficiency of this optical recalibration system depends on the number of points of the artificial stars projection.

For an ITT F 4012 RP tube used on a 14 mm photocathode diameter, figure 5 shows the difference in microns, between the position of real star and the position measured at one standard deviation versus the number of artificial stars of the optical recalibration system.

For the SED 04 star tracker developed for the SPACELAB's IPS which uses a set of 20 x 20 reference pattern points, the field/accuracy ratio, at 3 standard deviations, is $2.5 \cdot 10^4$, compatible with the specified requirements of : $2^\circ/0.75 \text{ arc sec} = 9600 (3\sigma)$ including all effects due to thermal drifts, chromatic distortions and magnetic field variations.

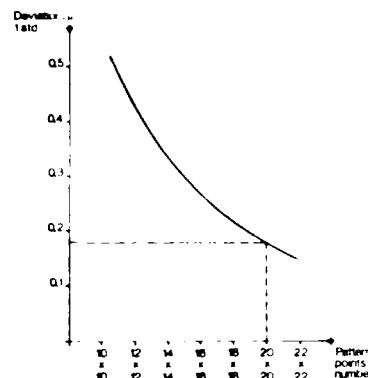


Figure 5. Recalibration accuracy according to pattern points number

Main characteristics and experimental results of the SED 04 star tracker developed for the Spacelab Instrument Pointing System

The main characteristics of the SED 04 tracker are summarized in table 1.

SED 04 Characteristics		
square field of view		$2^\circ \times 2^\circ$
optics	focal length	265 mm
	aperture diameter	80 mm
detector	image tube dissector	ITT 4012 RP
	photocathode	S 20
	instantaneous field of view	$200 \mu\text{m}/160 \text{ sec}$
electric	output data rate	$6 \times 17 \text{ bit/sec}$
	microprocessor operation frequency	16 MHz
	PROM memory	14 K bytes
	RAM memory	2 K bytes
	power (20 to 32 V)	21 W
size and mass		
optical head	- length	550 mm
	- diameter	210 mm
	- mass	7.5 kg
electronics box	- length	245 mm
	- width	210 mm
	- height	160 mm
	- mass	4.5 kg
* <u>Star mode</u>		
. MV visual magnitude		+ 2 to + 8
. degraded mode		+ 9
. colour temperature		3000 à 28000K
. sky background radiance		$8.4 \mu\text{lm}/\text{m}^2 \text{ sr}$
* <u>One-star mode</u>		
. maximum acquisition time		76 s
. tracking :		
- bandwidth		2.2 Hz
- noise equivalent angle		$0.5 \text{ sec}/\sqrt{\text{Hz}}$
* <u>Two-star mode</u>		
. acquisition time		152 s
. tracking :		
- bandwidth		1.1 Hz
- noise equivalent angle		$0.75 \text{ sec}/\sqrt{\text{Hz}}$
linearity and stability (at 3 standard deviations)		$0.75 \text{ sec}(3\sigma)$

Table 1 - Main characteristics of the SED 04 star tracker

The main experimental results are summarized in table 2. These results are representative of the measures carried out on the six SED 04 star trackers manufactured between 1979 and 1983.

SED 04 Characteristics	
One-star mode counting rate for a + 8 visual magnitude counting rate for a 8.4 $\mu\text{lm}/\text{m}^2$ sr sky back-ground . maximum acquisition time . noise equivalent angle for a $M_v = 8$ star and sky background	2800 c/s 500 c/s 70 s 0.45 sec/ $\sqrt{\text{Hz}}$
Two-star mode . + 8 visual magnitude . maximum acquisition time . noise equivalent angle for a $M_v = 8$ star and sky background	2800 c/s 132 s 0.7 sec/ $\sqrt{\text{Hz}}$
Stability . response linearity related to position in the field of view for 28°C and one colour temperature in the (3000-28000 K) range . deviation due to different colour temperature in the (3000-38000 K) range . deviation due to temperature variations in the (28 - 32°C) range for the optical head, and in the (+ 10 + 50°C) range for the electronic box, . magnetic field effect 0,4 gauss < 3 Hz Total stability (root sum square)	0.45 sec 3 σ 0.2 sec 3 σ 0.1 sec 3 σ 0.1 sec 3 σ 0.52 sec 3 σ
Field/accuracy ratio on each axis (3 standard deviations)	13850

Table 2 - Experimental results of the SED 04 star tracker

EVALUATION OF COMPARED ADVANTAGES AND DRAWBACKS OF IMAGE DISSECTOR TUBE AND CCD TYPE SOLID STATE DETECTOR STAR TRACKERS

The main parameters characterizing star tracker performances are :

- noise equivalent angular resolution for the dimmest star in terms of measurements rate,
- acquisition time for the dimmest star,
- field accuracy ratio.

Comparative assessment for these parameters is carried out starting from the IIT F4012 image dissector tube equipped with an S.20 photocathode and a RCA SID 52501 solid state detector.

Noise equivalent angular resolution

For an image dissector tube, the noise equivalent angular resolution (NA) is given by :

$$(\text{NA})^2 = \frac{1}{\text{IDF}} \cdot \frac{w^2}{\left(\frac{2}{B}\right)^2}$$

where w is the angular distance corresponding to the linear part of the character loop in arc sec,

and $\left(\frac{S}{B}\right)^2 = \frac{N_e T}{1 + 2 \frac{N_f}{N_e}}$ where T is the elementary tracking loop duration in seconds,
 N_e the star counting rate,
 N_f the sky background counting rate,

For a C circular measurement field, in degrees, a N number of optical aperture, a 14 mm photocathode diameter, a measurement spectral bandwidth corresponding to a S.20 photocathode, a 40 μm W angular distance, a 75 μm instantaneous field of view, a L_f sky background radiance in candela/m² and a star with a 6000 K colour temperature, the noise equivalent angular resolution in arc/sec is given by :

$$(RA)_{IDT}^2 = \frac{1}{T} \cdot 1,9 \cdot 10^{-7} C^4 N^2 2,51^{mv} [1 + 5,5 \cdot 10^{-3} C^2 L_f 2,51^{mv}]$$

For a CCD type solid state detector operated in tracking mode on a 3 x 3 pixels square elementary area using a circular image spot with a 2-pixel diameter, noise equivalent angular resolution (RA) is given by :

$$(RA)_{CCD}^2 = (\text{angular dimension of one pixel})^2 \cdot \frac{1}{\left(\frac{S}{B}\right)^2}$$

where pixel dimension is given in arc sec.

Without taking into account the CCD's reading noise, signal/noise ratio is given by :

where T is the elementary integration time chosen identical to the time of an elementary tracking loop for the dissector tube

$$\left(\frac{S}{B}\right)^2 = \frac{N_e T}{1 + 9 \frac{N_{ops}}{N_e} + 9 \frac{N_f}{N_e}}$$

N_e is the star signal in electrons/s
 N_{ops} is the dark current in electrons/s
 N_f is the signal due to the sky background on a pixel in electrons/s.

For a square measurement field, with a Co diagonal given in degrees, a No optical aperture number, a 30 μm square pixel, a 0.65 μm to 0.9 μm measurement spectral bandwidth, a 6000 K colour temperature star, a L_f sky background radiance in candela/m², a $N_{ops} = 12000 \cdot 2^{u/7}$, u being the CCD's operating temperature, the noise equivalent angular resolution : given by :

$$(RA)_{CCD}^2 = \frac{1}{T} \cdot 0,29 \cdot 10^{-7} Co^4 No^2 2,51^{mv} [1 + 6,5 \cdot 10^{-3} Co^2 L_f 2,51^{mv} + 4 \cdot 10^{-5} \cdot 2^{u/7} \cdot Co^2 No^2 2,51^{mv}]$$

Compared angular resolution for zero sky background level : $L_f = 0$

The R ratio = $\frac{(RA)_{CCD}^2}{(RA)_{IDT}^2}$ is given by :

$$R = 0,15 \left(\frac{No}{N}\right)^2 \cdot \left(\frac{Co}{C}\right)^4 [1 + 4 \cdot 10^{-5} \cdot 2^{u/7} \cdot Co^2 No^2 2,51^{mv}]$$

In the visual magnitude range covering $M_v = 1$ to $M_v = 8$, the field/magnitude relation may be approached by :

$$M_v = 10,2 - 4,1 \log C \quad \text{where C is the circular field in degrees}$$

which makes us sure that there is at least one star in the C field of view whatever is the orientation in the celestial sphere.

Comparison for these two detectors has to be made for equivalent solid angles in order to respect the field/magnitude relation. The $Co/\sqrt{2} \times Co/\sqrt{2}$ solid state detector square field has to be equivalent to the C image dissector circular field ; the C and Co fields are thus related by :

$$\left(\frac{Co}{C}\right)^2 = \frac{\pi}{2}$$

In order to keep in a realistic optical definition range for fields of view from 2° to 12° , we suppose the field and aperture being related by :

$$C.N = \text{constant} = 14$$

The $\frac{N_D}{N}$ ratio is considered equal to 1.4, so as to take into account that in its principle, the dissector tube is very little affected by an image spot distortion in the field. In this condition, the R ratio becomes :

$$R = 0,15 \cdot (1,4)^2 \cdot \frac{\pi}{4} \left[1 + 4 \cdot 10^{-5} \cdot 2,2^{10,2} \cdot (14)^2 \cdot 2,51^{10,2} \right]$$

In terms of measurement field, R becomes :

$$R = 0,74 \left[1 + 0,78 \cdot 10^{-2} \cdot 2,2^{10,2} \cdot 2,51^{10,2} - 4,1 \log e \right]$$

The curves shown on figure 6 clear the evolution of the R ratio, in terms of field of view and CCD's operating temperature. It shows that getting angular resolutions comparable for the 2 types of detectors requires to cool the CCD down to less than -20°C and that this cooling constraint is all the less severe as the tracker field of view increases. Nevertheless, as soon as the CCD's operating temperature is more than -20°C , angular resolutions for the 2 types of detectors are not comparable any more because the R ratio increases very quickly versus temperature.

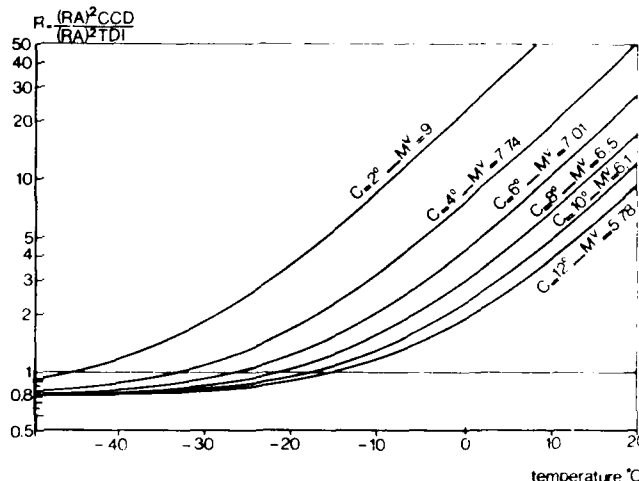


Figure 6. Compared angular resolution of the CCD and IDT in terms of field of view and CCD's temperature

Compared angular resolution for a high sky background level

In order to separate parameters, we suppose the CCD being cooled down to -40°C so as to neglect dark current influence.

Then R ratio is written :

$$R = 0,15 \cdot \frac{1 + 6,5 \cdot 10^{-3} \cdot C^2 \cdot 2,51^{10,2}}{1 + 6,5 \cdot 10^{-3} \cdot M^2 \cdot 2,51^{10,2}} \approx 0,74 \cdot \frac{6,5}{5,5} \cdot \left(\frac{C}{M} \right)^2 = 1,4$$

As we have seen that for high sky background radiances, the two types of detectors are comparable :

Acquisition time

As we have seen, the first acquisition phase includes detection and identification of the signal, which the signal exceeds a given threshold (the threshold being chosen in order to ensure that the threshold noise exceeding a given value) and the identification of the signal (the threshold noise exceeding a given value) in order to fulfill the condition :

$$S > N$$

For a two $\sigma = 90\%$, allowing to neglect dark current, the acquisition time will be lower than

the elementary acquisition time for a given ratio in the same order of magnitude

Considering ϕ_c as the total field of view diameter and ϕ_{cvi} the instantaneous field of view diameter of the dissector tube :

$$T_{\text{acquisitionIDT}} = \frac{\phi_c^2}{2 \phi_{cvi}} \cdot T_{\text{elementary}}$$

$$T_{\text{acquisitionCCD}} = T_{\text{elementary}}$$

It has to be noticed that for most of high accuracy applications, the tracking field may be limited because rough orientation in the celestial sphere is known and thus may be used in order to limit to more than necessary the number of searching steps for an image dissector tracker.

Field-accuracy ratio

Theoretical and experimental results show that when using a 400 points artificial stars projector on an image dissector tube photocathode, the resulting field/accuracy ratio is around $2.5 \cdot 10^4$. Using the optical recalibration system allows to break free from energy distribution variations in the image spot with respect to position in the field and size variations of the image spot produced by a defocusing related to photocathode-optics relative axial shifting. Indeed, these fluctuations are considered by means of pattern artificial stars measurement because in a reduced field limited by a square of 4 successive pattern points, residual variations have to be taken into account. Only the angular resolution is a bit affected because it depends on the transfer function slope of the tracking loop depending linearly on the image spot size.

The $2.5 \cdot 10^4$ field/accuracy ratio is not a limit, it may be bettered either by increasing the number of artificial stars of the optical recalibration system, or by using an additional correction algorithm taking into account local geometric distortions related to non-linearity of deviation coils.

A $5 \cdot 10^4$ to $8 \cdot 10^4$ field/accuracy ratio objective is quite foreseeable on operational equipment

For a solid state detector, the use of an interpolation process is required. It leads to dispatch signal flow on several pixels and to use the responses given by single pixels in order to determine the geometric position of image spot in pixel fraction. Limitations brought by pixel non-uniformity response, pixels position and geometric size deviations, energy dispatching fluctuations in the image spot in terms of position in the field and radiance spectral composition lead to limit the interpolation quality of about 1/30th to 1/100th pixel in terms of treatment complexity and calibration (pixel response modelisation, average sensitivity memorization for every pixel, geometric defects corrections by using a linear regression...). This leads to a 7500-30000-range field/accuracy ratio on the theoretical level with the existing CCDs.

Reaching these performances requires firstly to cool down the solid state detector and secondly either to use corrections taking into account detector physical parameters (dark current and sensitivity of all pixels), or to use a component for which the variations between those physical parameters are lower than 1 % for the whole detector. Positioning stability relative to the solid state detector and associated optics is very critical too. A transversal relative displacement of $\pm 0.5 \mu\text{m}$ corresponds to an alignment defect equivalent to 1/30th pixel which can only be corrected by using an optical recalibration system with at least three artificial stars. Axial positioning variations introduce energy distribution variations in the image spot ; they are the more critical as the optics has a low aperture number.

All these energy distribution constraints in the image spot, of thermo-mechanical conception and calibration parameters evolution (dark current, sensitivity of all pixels) in ageing and under radiance constraints in space environment do not easily fit requirements of an operational equipment ; this leads to think that for a simple structure operational equipment, the reasonable limit for interpolation capacity is better than 1/10th pixel, thus a field/accuracy ratio of around 2000 to 4000.

In order to improve the field/accuracy parameter, we need on one hand, to use an optical recalibration system, on the other hand, to develop image spot defocusing methods allowing the equipment not to be sensitive to variations in the field of view and to energy distribution variations in the spot for optics-detector axial relative displacements. A well adapted - so it seems - solution consists in using an electronic defocusing induced by an image intensifier tube allowing in other respects, to use a low aperture optics so as to solve energy distribution problems in the image spot which are only related to optics. Relative sensitivity homogeneity constraint of pixels (better than 1 %) remains a leading parameter for interpolation quality ; a solution consisting in using a solid state detector illuminated from the backside should allow to improve homogeneity quality of response on the whole detector.

This type of concept, related to the use of a low aperture optics, introduces signal/noise reduction and requires to develop photocathodes with a quantum efficiency and a sensitivity improved in the 600-900 nm bandwidth (for instance, AsGa photocathodes) ; once developed, this type of photocathode could be integrated in an image dissector tube.

An approach solution for a high performance CCD solid state detector star tracker leads to the following structure :

- solid state detector, with a response homogeneity better than 1 %, illuminated from the backside and thermo-controlled around 0°C (profit contribution of intensifier tube allow to limit dark current effect by reducing integration time),

- solid state detector-intensifier tube coupling,

- low aperture optics including a 3 artificial stars optical recalibration system,

which objective performances in space environment would be around 1/50th pixel, thus a 10000 to 20000 field/accuracy ratio.

REFERENCES GYROSCOPIQUES POUR APPLICATIONS SPATIALES

par

J. Resseguier et G. Monier
SAGEM
B.P. 51
95612 Cergy Pontoise Cedex
France

Dans cette communication on présente d'abord les gyroscopes SAGEM disponibles pour les applications spatiales en montrant l'évolution technologique entre:

- Les gyroscopes flottants;
- Les gyroscopes sec accordés.

Dans chaque type, les caractéristiques essentielles sont commentées:

- Fiabilité (MTBF)
- Stabilité à long terme
- Bruit
- Consommation.

On déduit ensuite une étude de redondance théorique et expérimentale destinée à optimiser l'organisation d'un système gyroscopique à senseurs multiples.

Enfin est présenté un équipement de mesures d'attitudes:

- à composants liés
modulaire conçu pour équiper un satellite.

FAULT TOLERANT SPACEBORNE COMPUTER

by
 Davis R. Bates, Jr.
 Raytheon Company
 Equipment Division
 528 Boston Post Road
 Sudbury, Massachusetts
 01776
 U.S.A.

James E. Schell
 Director
 Center For Tactical Computer Systems
 U.S. Army Communications - Electronics Command
 Fort Monmouth, New Jersey 07703
 U.S.A.

Dr. G. Robert Couranz
 Raytheon Company
 Equipment Division
 528 Boston Post Road
 Sudbury, Massachusetts
 01776
 U.S.A.

SUMMARY

Autonomous spacecraft control and data processing have recently become realizable due to advanced fault tolerant techniques coupled with the development of very high scale integrated circuits. A method of utilizing present integrated circuit componentry with future technology insertion has been developed on the FTSC Program. The FTSC is designed to support a variety of spaceborne missions with a 95% probability of surviving, unattended and without degradation in performance, for five to seven years. Besides easing the ground station scheduling requirements, this capacity increases the overall satellite survivability.

The purpose of this paper is to present a concept of achieving the survivability by the utilization of various fault tolerant techniques. These concepts of fault tolerance, as well as technology insertion, are also being incorporated in the Military Computer Family Program currently under development at the U.S. Army's CECOM Facility at Ft. Monmouth, New Jersey, U.S.A.

The Fault-Tolerant Spaceborne Computer (FTSC) is a general-purpose, high-reliability, radiation-hardened digital computer designed to meet the Air Force on-board computational and survivability requirements of the 1983-1993 time frame. The FTSC is designed to support a variety of spaceborne missions with a 95% probability of surviving, unattended and without degradation in performance, for five to seven years. It is capable of self-reconfiguration and resumption of computation following internal element failures, power transients, and radiation events. With these features, the FTSC provides autonomous spacecraft control via on-board programs, as opposed to present systems which require excessive ground interaction. Besides easing the ground station scheduling requirements, this capability increases the overall satellite survivability. The basic FTSC characteristics are provided in Table I. As the table shows, the FTSC also has outstanding performance characteristics as a computer. This permits very accurate attitude control, navigation, and guidance. Since it can be reprogrammed to accommodate changing mission requirements, the FTSC offers a high-degree of on-orbit flexibility and reduces the probability of early obsolescence.

A system block diagram of the FTSC (Figure 1) shows the computer is segregated into a number of functional elements, with each element protected by at least one spare. In addition, several of these elements, notably the memory elements and the address and data buses, are provided with redundancy at the sub-element level through the use of "Ripple" switches. This concept greatly reduces the redundant hardware that would otherwise be required. The Configuration Control Unit (which is the ultimate guardian of the system) has been reduced in complexity by distributing much of its function to other FTSC elements. High fault-coverage is assured through extensive use of coding, and single-point failure mechanisms have been avoided by, among other things, providing at least two paths for all critical signals. A short description of each of the functional elements follows.

CENTRAL PROCESSING UNIT (CPU)

The CPU implements all arithmetic and logical operations required to process application and fault recovery programs. There are four CPU modules in the FTSC, two of which are enabled during normal operation. (Note: The FTSC may accommodate six CPUs for longer life.) One of the enabled CPUs is designated as the active module, the other as the monitor. Each CPU has the following features:

- Executes the 112 instructions as defined in Table IV. These instructions are performed on 32-bit data words and include integer, floating point, and logical operations.
- Nine addressing modes:
 - Register-Register
 - Immediate
 - Upper Immediate
 - Direct
 - Indirect
 - Indexed, Post-Increment
 - Indexed, Pre-Decrement
 - Indexed
 - Indexed, Indirect
- Microprogrammed using an internal read-only memory (ROM)
- Eight program addressable 32-bit general purpose registers, plus an extension register, a program counter, status registers, a memory address register, a memory data register, and four working registers.
- Capable of recognizing and processing ten internal levels of interrupts.
- Incorporates a bus arbiter to resolve bus access conflicts among the CPU, SIU and DMA.
- Capable of operating in either of two modes: active or monitor. The active CPU provides all information and control to the system; the monitor CPU compares the outputs from the active CPU with those it would have generated if it were the active CPU and reports any discrepancy to the Configuration Control Unit (CCU).
- A 2K x 32-bit reconfiguration read-only memory (ROM) which contains the reconfiguration program. This program is initiated in response to a fault interrupt.

CONFIGURATION CONTROL UNIT (CCU)

The CCU has the capability to interrupt the CPU whenever a fault is detected and to provide sufficient status information to ensure a successful reconfiguration. There are three CCU modules in the FTSC which operate in a triple-modular-redundant (TMR) configuration. Each CCU performs the following functions:

- Causes the fault recovery program to be initiated by issuing a fault interrupt to the CPU whenever a failure is detected by one or more of the fault detection monitors. (The fault interrupt is not issued if the specific failure detected has been previously masked under program control.)
- Changes the CPU, address bus, and/or data bus configurations whenever a failure indicates that such a change is necessary to effect a successful recovery.
- Accepts program flag inputs from the CPU. The functions performed by these flags include masking certain types of faults for testing and for causing the CPU, address bus, or data bus states to be changed.

MEMORY MODULE (MM)

The main memory in the FTSC is composed of a number of independent semi-conductor memory modules each with a capacity of 4096 words. The design permits expansion to 24 memory modules, 15 of which can be active, directly addressable memory (61,440 words). The remaining memory modules are spares. Each memory module provides:

- Random access storage for FTSC programs and data.
- A word size of 41 bits, consisting of 38 active bits and three spare bits. The 38 active bits consist of 32 bits of data and six bits of code.
- The capability to select the active and spare bit-line positions under program control. This selection applies to all words in the module so that if a bit is identified as a spare, that bit is a spare in each word of the module.
- Assignment of any one of the following 16 sets of continuous addresses: 0 to 4E-1, 4E to 8F-1, ..., 59E to 59F-1, 5F indicates 1024 words.

- The ability to individually write-protect each memory subblock under program control. (Each memory module is divided into four subblocks of 1024 contiguous locations each.) Any attempt to store into a write-protected area leaves the information in that address unchanged and causes a fault.
- The capability of detecting multiple bit errors using the 6-bit code in each word. The memory can also correct single bit errors when put in the error correcting mode under program control.
- Activation or deactivation under program control. (Following activation, a memory module does not respond to read or write operations until so directed under program control.)

SERIAL DATA BUS

The serial data bus subsystem allows up to 62 external interfaces to the system. The serial bus is controlled by the Serial Interface Unit (SIU) and is connected to the peripherals through dual-redundant Device Interface Units (DIU). The bus is a uni-directional, daisy-chain system with dual-redundant serial links between each contiguous pair of DIUs along the bus.

SERIAL INTERFACE UNIT (SIU)

The FTSC contains two SIU modules connected in a dual-redundant configuration. Only one SIU is active at a time; it can be deactivated and the other activated under program control when a failure is detected. Each SIU module has the following capabilities:

- Transfer blocks of words between the serial bus and main memory using a direct-memory-access capability.
- Sequentially access any number of locations in a program specified block (module) of memory. It fetches a word from each address, transmits the word over the serial bus, and stores the returned (possibly modified) word back into the memory location from which it was fetched.
- Synchronize the initiation of a block transfer with the real-time interrupt (RTI). (In the Immediate Mode, the block transfer is initiated with the next transmitted word.) A new block transfer cannot be initiated if the SIU is already in the process of transmitting a block.
- Generate two levels of interrupts: an end-of-block interrupt signalling the completion of the current block transfer, and a high-priority interrupt indicating peripheral initiated interrupts or a failure condition.

DEVICE INTERFACE UNIT (DIU)

Low speed user devices are connected to the serial bus through dual-redundant pairs of DIUs. Both DIUs in each pair are powered at all times by the peripheral in which they are located. Each DIU has the following characteristics.

- Transmit or receive status (command) information or data in accordance with the OP CODE field of only those received words containing its unique pin programmed address.
- Following a peripheral-initiated interrupt, the module modifies the first available serial bus word to indicate an interrupt to the SIU (unless a higher priority interrupt is already indicated in the word). The DIU repeats the interrupt every third real-time interrupt (RTI) until it is acknowledged.

DIRECT MEMORY ACCESS (DMA)

The DMA provides a high speed parallel interface to the FTSC. There are two dual-redundant pairs of DMA modules in the FTSC. Only one DMA of the pair is activated, while the other acts as a nonoperating or standby spare. Either can be activated or deactivated under program control. Each DMA module is capable of the following:

- Transferring blocks of information, without CPU intervention, directly from any active memory module to the peripheral or from the peripheral to memory.
- Passes on commands sent by the CPU to control the peripheral.
- Generates two levels of interrupts: an end-of-block interrupt signalling the completion of the current block transfer, and a high-priority interrupt indicating a peripheral-initiated interrupt or a failure condition.

TIMING UNIT (TU)

The Timing Unit consists of two powered Timing Modules in a dual-redundant configuration. One module is active, while the other serves as a standby spare. Each Timing Module has the following features:

- When active, it generates a system clock at a frequency of 2.5 MHz.
- When active, it generates a real-time interrupt (RTI) every 4.4032 ms. Capability is provided to decrease this period by a factor of 2 or to increase it by a factor of 2 or 4.
- Provides dual status signals to other FTSC modules to indicate its active/standby state.
- Contains dual monitors to check its system clock and internal real-time interrupt outputs. If an out-of-tolerance condition is detected, the defective Timing Module enters the standby state and the other Timing Module takes over the active role.

POWER UNIT (PU)

The Power Unit consists of two Power Modules connected in a dual-redundant configuration. One module is active while the other serves as a standby spare. Each Power Module converts the primary power input to the internal levels required. Each Power Module has the following capabilities:

- When active it provides the internal voltages required by the FTSC.
- Provides dual status signals to the other FTSC modules to indicate its active/standby state.
- Contains dual monitors to examine its output voltage levels. If an out-of-tolerance condition is detected, the defective Power Module enters the standby state and the other Power Module takes over the active role.

CIRCUMVENTION UNIT (CU)

The Circumvention Unit is designed to protect the FTSC during radiation events and power interruptions by clamping all memory, DMA, and SIU modules. There are three Circumvention Modules in the FTSC operating in a triple-module-redundant (TMR) configuration. Each Circumvention Module has the following characteristics:

- Can accept inputs from distributed radiation detectors.
- Monitors for the status signals indicating out-of-tolerance power conditions.
- During the period following detection of radiation or power out-of-tolerance conditions, generates clamp signals which inhibit all I/O and prevent memory contents from being changed.

HARDENED TIME (HT)

The Hardened Timer keeps track of real time during radiation events, power module failures, timing module failures, and reconfiguration. There are three Hardened Timer Modules in the FTSC operating in a triple-modular-redundant (TMR) configuration. Each Hardened Timer Module is capable of the following:

- Being activated upon the occurrence of a radiation event, a power failure, a timing failure, or the detection of any unmasked fault in the FTSC.
- Accumulating the number of real-time interrupts (RTI) which occur during the time in which the timer is active.
- Reading and resetting the accumulated time contents under program control.
- Accumulating elapsed time.

INSTRUCTION FORMAT

The FTSC instruction format consists of five fields: OP CODE, RB, RA, AM and ADDRESS. The OP CODE field identifies the instruction to be executed. The RB field specifies one of the eight general-purpose registers used as an operand or destination register. The RA field specifies one of the same set of eight general-purpose registers used as an address index, operand or destination register. The AM field specifies the address mode to be used in defining the effective address for the instruction. Table II defines the operation of each address mode and the calculation of ea for each mode. Two operands (OP1 and OP2) have been defined to identify the difference among the modes for fetch-type instructions. For fetch-type instructions in which the content of r_p is not specified to be modified, the content of r_p will be replaced by OP2. The ADDRESS field is used in defining the ea for modes 1-7 and the address specified is designated "a" in the address mode table.

OP CODE	RB	RA	AM	ADDRESS
0	6	9 10	12 13	15 16
				31

FIELD	BITS	MEANING
OP CODE	0-6	Operation code
RB	7-9	Specifies one of 8 operand/ destination registers
RA	10-12	Specifies one of 8 index/ operand/destination registers
AM	13-15	Specifies address mode
ADDRESS	16-31	Specifies operand or address

ADDRESS MODES

Address Mode	OP CODE MSB	Description	Instruc- tion Type	Operand Rules*
0	0	Register- Register	Fetch	OP1 = (r_r) = (ea) OP2 = (r_o)
1	0	Immediate	Fetch	OP1 = sign extended a = (ea) OP2 = (r_o)
1	1	Upper- Immediate	Fetch	OP1 = $a \times 2^{16}$ = (ea) OP2 = (r_o)
2	0	Direct	Fetch	OP1 = (a) = (ea) OP2 = (r_o)
3	0	Indirect	Fetch	OP1 = ((a)) = (ea) OP2 = (r_o)
4	0	Indexed, Post- increment	Fetch	OP1 = (a + (r_r)) = (ea) OP2 = (r_o) (r_r) ← (r_r) + 1
5	0	Indexed, Pre- decrement	Fetch	(r_r) ← (r_r) - 1 OP1 = (a + (r_r)) = (ea) OP2 = (r_o)
6	0	Indexed	Fetch	OP1 = (a + (r_r)) = (ea) OP2 = (r_o)
7	0	Indexed, Indirect	Fetch	OP1 = ((a + (r_r))) = (ea) OP2 = (r_o)
0	1	Register- Register	Store	r_r = ea
2	1	Direct	Store	a = ea
3	1	Indirect	Store	(a) = ea
4	1	Indexed, Post- increment	Store	a + (r_r) = ea (r_r) ← (r_r) + 1
5	1	Indexed, Pre- decrement	Store	(r_r) ← (r_r) - 1 a + (r_r) = ea
6	1	Indexed	Store	a + (r_r) = ea
7	1	Indexed, Indirect	Store	(a + (r_r)) = ea

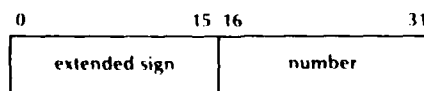
* (X) = the content of X.

Table II

DATA FORMATS

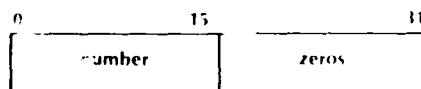
The logical data, integer data, floating-point data, immediate data, upper-immediate data, double-precision floating-point data and program status words are defined in this Figure.

Immediate Data (IM)



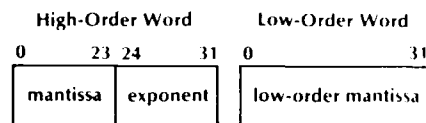
Immediate numbers are treated as 16-bit integers. The sign bit occupies bit 16 and is extended by immediate mode instructions to the upper 16 bits (0-15) of the word before the value is used in the computation. Immediate numbers may be used as logical or integer data.

Upper-Immediate Data (UIM)



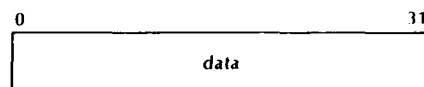
Upper-immediate numbers are treated as 16-bit, 2's complement integers. The value of the instruction address field is multiplied by 2^{16} before the value is used in the computation. Upper-immediate numbers may be used as logical, integer or floating-point data.

Double-Precision Floating-Point Data (DPF)



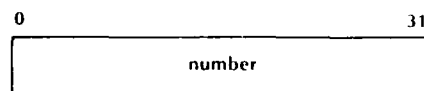
The high-order word of double-precision floating-point numbers has a 24-bit 2's complement normalized, fractional mantissa and an 8-bit 2's complement exponent. The low-order word is a 32-bit continuation of the high-order mantissa.

Logical Data (L)



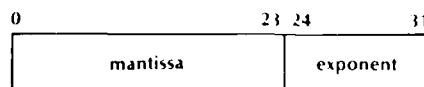
All bits are treated identically

Integer Data (I)



Integers are represented in a 32-bit, 2's complement form

Floating-Point Data (F)



Floating-point numbers have 24-bit, 2's complement, normalized fractional mantissa and an 8-bit, 2's complement exponents.

A floating-point zero is represented by an all zero mantissa and an 80 (hex) exponent.

INSTRUCTION SET

The FTSC has a comprehensive and flexible instruction set which contains both integer and floating-point computational instructions.

The instruction algorithms, and the central-processing unit hardware, are designed to provide maximum execution speed for a computer which has very low average electrical power consumption.

Among the instructions in the repertoire are many that substantially reduce the execution time of application programs which are typical of spacecraft and satellite missions. The instruction format allows for up to 128 instructions, of which 112 have been specified for the FTSC. This, along with 9-address modes, provides a high degree of flexibility for the programmer.

As the computer is primarily intended for spaceborne guidance, navigation and satellite management functions, special emphasis has been made to provide a complete set of arithmetic operations. Both integer fixed-point and fractional floating-point instructions are provided. Additional speed is derived from double-precision floating-point operations and vector operations. The floating-point vector operations include the powerful Inner Product and Scalar Multiply as well as the normal arithmetic functions of vector addition and vector subtraction. An extensive set of conditional jump and skip operations combined with expanded immediate mode addressing provides for efficiencies in program storage needs. Many spacecraft functions have a need for controller type operations, where data matching and bit inspection is required. Logical shift and mask instructions are provided to satisfy these requirements.

The first instructions are divided into two major categories: instructions - those which are generally available to the application programmer and system instructions - those which are restricted to application use. Of the 112 instructions, 94 are applications instructions and the other 18 fall into the system category.

Some typical instruction execution times are presented in Table III. A complete list of 112 instructions with their names is Table IV.

Table III
TYPICAL
INSTRUCTION EXECUTION TIMES

INSTRUCTION	EXECUTION TIME
Direct Mode (2)	
(STR) STORE REGISTER	2.8/5
(LDR) LOAD REGISTER	2.8
(JMP) JUMP	3.2
(AND) AND LOGICAL	2.8
(ADD) ADD INTEGER	2.8
(ADDF) ADD FLOATING	8.0
(MPY) MULTIPLY INTEGER	11.6
(MPYF) MULTIPLY FLOATING	10.4
(DIV) SHORT DIVIDEND DIVIDE INTEGER	16.0
(DIVF) DIVIDE FLOATING	17.2
(SRTF) SQUARE ROOT FLOATING	52.0
(VIPF) VECTOR INNER PRODUCT FLOATING	46.4
(ARL) ARITHMETIC LONG SHIFT (10 BITS RIGHT)	7.6
Register-Register Mode (0)	
(LDR) LOAD REGISTER	1.6
(ADD) ADD INTEGER	1.6
(ADDF) ADD FLOATING	6.8

Table IV
FTSC INSTRUCTION SET

LOAD/STORE

Op Code	Mnemonic	Name
00	LDR	Load Register
02	LDR2	
03	LDR3	Load Multiple Registers
04	LDR7	
26	LDS2	
2F	LDS3	Load Scalar into Registers
3A	LDS7	
01	LDE	Load Extension Register
30 E*	LW0	Load Working Register W0
31 E*	LW1	Load Working Register W1
32 E*	LW2	Load Working Register W2
33 E*	LW3	Load Working Register W3
0A E*	LA0	Load Active Only
0B E*	LMO	Load Monitor Only
09	LDC	Load One's Complement
05	LDN	Load Negative Integer
06	LDNF	Load Negative Floating
07	LDA	Load Absolute Value Integer
08	LDAF	Load Absolute Value Floating
3B	MOV	Move
40	STR	Store Register
62	STR2	
63	STR3	Store Multiple
68	STR7	
42	STD	Store Duplex
44	STD2	
45	STD3	Store Multiple Duplex
46	STD7	
47	STZ	Store Zero
41	SZD	Store Zero Duplex
41	STE	Store Extension Register
48	SPS	Store PSW
4A	SPC	Store PSW Duplex
6D	SWP	Swap
6E	INC	Increment
6F	INCD	Increment Duplex
KA E*	SW0	Store Working Register W0
KB E*	SW1	Store Working Register W1
K3 E*	SW2	Store Working Register W2
KB E*	SW3	Store Working Register W3
49 E*	STH	Store to Half Address
4B E*	SBPA*	Store with Bad Address Parity 0-15
4C E*	SBPA0	Store with Bad Address Parity 0-15
4D E*	SRPD*	Store with Bad Data Parity 0-15
4E E*	SRPD0	Store with Bad Data Parity 0-15

JUMP

50	JMP	Jump
51	JM	Jump if Negative
58	JM2	Jump if Negative or Zero Integer
52	ZF	Jump if Zero Integer
54	NZ	Jump if Not Zero Integer
4F	DZ	Jump if Discrepant Zero
56	PC	Jump if Positive and Not Zero Integer
59	NM2	Jump if Negative or Zero Floating
53	ZF	Jump if Zero Floating
55	NZ	Jump if Not Zero Floating
57	PC	Jump if Positive and Not Zero Floating
5A	PC	Increment by Jump if Not Zero Integer Integer
5C	PC	Jump if Discrepant Not Boolean Integer
5D	PC	Jump if Discrepant Boolean Integer
5E	SR	Jump if Signed
5F	PC	Increment by Jump if Signed Integer Integer
59	PC	Jump if Signed

TEST AND SKIP

Op Code	Mnemonic	Name
2C	ASZ	AND Skip if All Zeros
2A	ASNZ	AND Skip if Not All Zeros
2D	OSZ	OR Skip if All Ones
29	OSN	OR Skip if Not All Ones
35	CSZ	Compare Skip if Equal
34	CONE	Compare Skip if Not Equal
36	CSZ	Compare Skip if Greater Integer
37	CSZ	Compare Skip if Less Integer
38	CSZF	Compare Skip if Greater Floating
39	CSZF	Compare Skip if Less Floating

INTEGER ARITHMETIC

19	ADD	Add Integer
1A	SUB	Subtract Integer
1B	MPY	Multiply Integer
1C	DIV	Short Dividend Divide Integer
1D	DIVF	Long Dividend Divide Integer
2R	ADD	Add Integer With Carry
2F	CF	Carry Integer to Floating

FLOATING POINT ARITHMETIC

19	ADD	Add Floating to Integer
1A	SUB	Subtract Floating
1B	MPY	Multiply Floating
1C	DIV	Divide Floating
1D	ADD	Add to R of Floating
1E	ADD	Add with Floating
1F	ADD	Add to R of Floating
20	ADD	Add to M of Floating
21	ADD	Add to R of Floating
22	ADD	Add to M of Floating
23	ADD	Add to R of Floating
24	ADD	Add to M of Floating
25	ADD	Add to R of Floating
26	ADD	Add to M of Floating
27	ADD	Add to R of Floating
28	ADD	Add to M of Floating

LOGICAL

19	AND	AND Integer
1A	AND	AND Integer
1B	AND	AND Integer
1C	AND	AND Integer

SHIFT/ROTATE

24	ARL	Arithmetic Right Shift
25	ARL	Arithmetic Right Shift
26	ARL	Arithmetic Right Shift
27	ARL	Arithmetic Right Shift
28	ARL	Arithmetic Right Shift

MISCELLANEOUS

19	PC	Increment by 1
1A	PC	Increment by 2
1B	PC	Increment by 4
1C	PC	Increment by 8
1D	PC	Increment by 16
1E	PC	Increment by 32
1F	PC	Increment by 64
20	PC	Increment by 128

SOFTWARE

The FTSC has a software package for the development of software which is matched by few computers in its class. The package contains three compilers, an assembler system, a simulator, debugging aids, diagnostic programs, a multi-tasking executive, input/output handlers and a mathematical subroutine package.

Compilers

- JOVIAL J73/1
- FORTRAN
- SPL

All compilers output symbolic assembly language and are enhanced with FTSC required functions such as rollback insertion and duplex data specification constructs.

Assembler System

- Marco Preprocessor - A powerful and effective tool for generating assembly language programs.
- Assembler - Accepts FTSC symbolic assembly language and produces relocatable object text and source listings with features such as External Symbols list, Entry Point Symbols list, Cross Reference list, Instructions Used list and Segment Description maps.
- Linkage Editor - Links together separately assembled relocatable outputs produced by the assembler and produces a "load module." Features extensive inputs analysis and utilizes by option either first fit or best fit algorithms for efficient memory management. Outputs detailed memory utilization maps.

Simulator System

- FTSSIM - An interpretive FTSC simulator with advanced debug capability. It includes CPU functions, CCU functions, DMA functions and memory functions. It is written in Pascal and is highly transportable.
- Loader-Program used to format the load modules produced by the linkage editor into the various formats required by the simulator and the FTSC hardware.

Debugging Aids

- Test Station - The Automated FTSC Test System (a PDP-11 system) communicates with CDC 6600/CYBER series computers for the transfer of FTSC loadable files. The ATIS is also interfaced with the FTSC on DMA, DIU and Control Panel ports. This interface allows great flexibility in loading and controlling FTSC programs as well as simulation of peripherals.
- Debug Package - FTSC resident program that provides such features as multiple breakpoints, memory dump, memory register inspection and modification, etc.

Diagnostics

- CPU Test - A comprehensive program for verifying the operation of the CPU and its system interfaces.
- Memory Test - A comprehensive program for verifying the operation of the memory system including all internal functions such as write protect and tripping.
- I/O Tests - Tests for verifying the operation of the DMA, DIU and DIU.

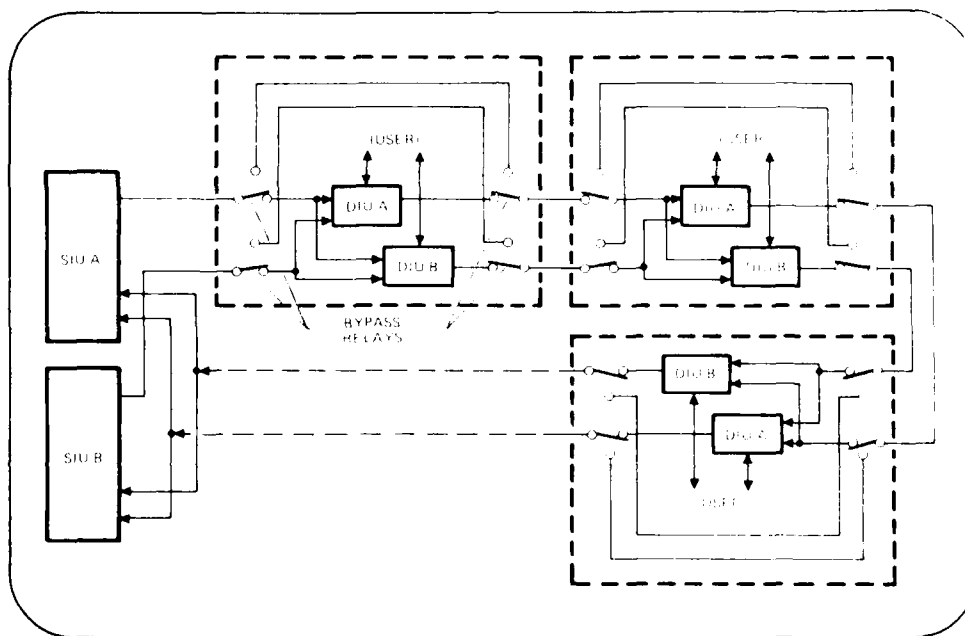
Operational Utilities

- Executive - Controls the multi-tasking executive. Features all the required services for all programs except I/O, error and program failure control which are handled by the I/O handlers.
- I/O handlers - Controls the input and output of data to and from the hardware. Includes a handler for the DIU and the DMA.
- Math Library - A collection of mathematical routines.
- Linker - A program which links the object files into a loadable module.

SOFTWARE SUMMARY

Software	Host Computer
JOVIAL J73/I	DEC 10
FORTTRAN	CDC 6000/CYBER SERIES
SPL	CDC 6000/CYBER SERIES
MACROPROCESSOR	CDC 6000/CYBER SERIES
ASSEMBLER SYSTEM	CDC 6000/CYBER SERIES
SIMULATOR SYSTEM	CDC 6000/CYBER SERIES
TEST SYSTEM	DEC PDP-11
DEBUG	FTSC
DIAGNOSTICS	FTSC
OPERATIONAL	FTSC

SERIAL DATA BUS SYSTEM



The Serial Data Bus (SDB) provides the FTSC with the capability to service up to 32 peripheral devices. Information is transmitted over the serial bus at at least a 325 Kbps basic data rate. The basic elements of the Serial Data Bus system are the dual-redundant Serial Interface Unit (SIU), which control bus communications, and the dual-redundant Device Interface Units (DIU) located at each of the peripheral interfaces. Only one of the two SIUs is activated at a time; while both of the DIUs at each interface are activated continuously.

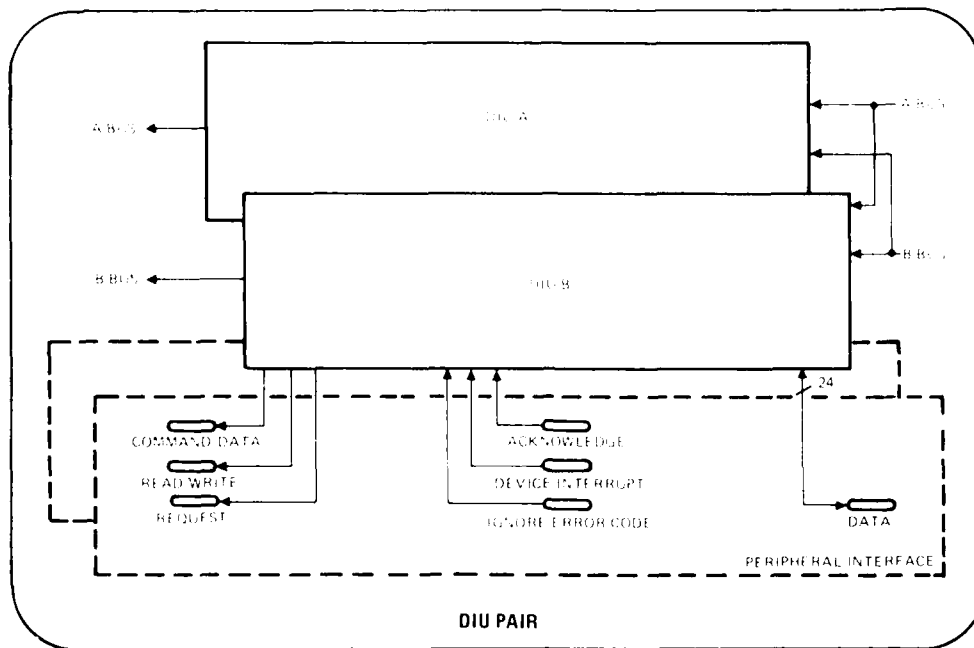
The SIU coordinates the transfer of blocks of words between the serial data bus and FTSC main memory by utilizing a direct memory access capability. The SIU is also capable of transferring serial data bus words on a single word basis utilizing an interrupt capability to notify the CPU of user requests.

The serial bus is a reconfigurable daisy-chain system with dual-redundant serial data links between each contiguous pair of DIUs along the bus, between the SIU and the first DIU and between the last DIU and the SIU. Transmission on the bus is uni-directional. Serial data originated at the SIU is received at each DIU along the bus where it is reconstructed or modified before being retransmitted to the serial bus.

The serial bus link interface over which data will be transmitted is optical fiber. The output driver has the capability of driving at least 60 feet of 75 ohm characteristic impedance twisted shielded pair cable.

The user must provide the connectors for the dual sets of cables both coming into the DIU pair from the Serial Bus and going from the DIU pair, back out to the bus.

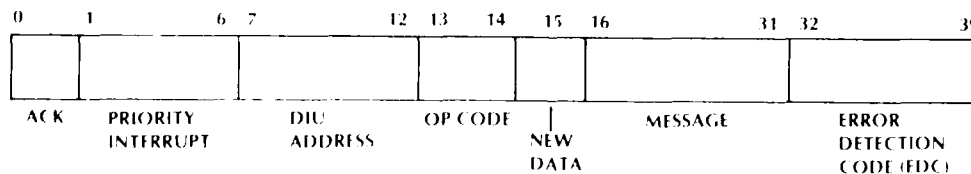
DEVICE INTERFACE UNIT (DIU)



Each of the dual DIUs provides a 24-bit parallel FTSC interface for the peripheral device with which it is associated. The peripheral must provide power and house each DIU of its associated pair. Each DIU monitors two serial data bus lines, although it actively processes data received on only one of the two. The DIU monitors the selected bus for faulty bus transmissions and switches to the alternate bus, if a fault condition is present on that bus. Relays are provided for by-passing a DIU pair should the associated peripheral lose power.

Each DIU transmits or receives status/command information or data as specified by the OP CODE field of the serial bus word. (The serial bus word format is shown below.) As the address field of each serial word is shifted into the DIU, each bit is tested to determine if it is consistent with the DIU's pin-programmed address. When the DIU detects its own address, it performs the following:

- If the OP CODE indicates receive message and the DIU is able to accept the message, the message is sent to the peripheral.



- If the OP CODE indicates receive command, the DIU checks the NEW DATA bit to see if the command is intended for the peripheral as well. If the peripheral is indicated, the COMMAND/DATA LINE is set high and the information is passed across the interface, and the DIU executes the portion of the command intended for it.
- If the OP CODE indicates transmit message, the DIU transmits the message, only if the message has been updated by the user since the last time it was transmitted.
- If the OP CODE indicates transmit status, the DIU transmits the status word. The first eight bits of the status word are reserved for DIU status and the second eight bits for peripheral status.

Data transfers between the DIU and its associated peripheral are effected via a 24-bit (24) bi-directional data lines, three (3) handshaking control lines and one (1) static control line. Three of the control lines come from the DIU: Request, Read/Write, and Command/Data and three control lines come from the peripheral: Interrupt, Acknowledge, and Ignore Error Code (static).

The request line is raised by the DIU to initiate a transfer to or from the peripheral. When this line is raised, the peripheral must act upon the current state of the other control lines.

The Read/Write line indicates the direction of the transfer over the DIU/Peripheral interface. The line is set to the high state when the transfer is from the peripheral to the DIU and the line is set low when information is to be transmitted from the DIU to the peripheral.

The Command/Data line enables the peripheral to distinguish between command/status information and data. If this line is in the high state, information transferred from the DIU to the peripheral, and to the DIU from the peripheral, is defined as command and status, respectively. If the line is at the low state, information transferred in either direction is defined as data.

The Acknowledge line is set high by the peripheral, in response to a DIU request, after the peripheral has either 1) gated data or status onto the data lines (Read/Write set high), or 2) latched the data or command present on the data lines (Read/Write set low).

For peripheral initiated transfers, the Interrupt line is set high by the peripheral. In response to the interrupt, the DIU set the Command/Data and Read/Write lines high and then raises the Request line. The peripheral must then raise the Acknowledge line and clear the Interrupt to indicate to the DIU that the device status has been gated onto the data lines.

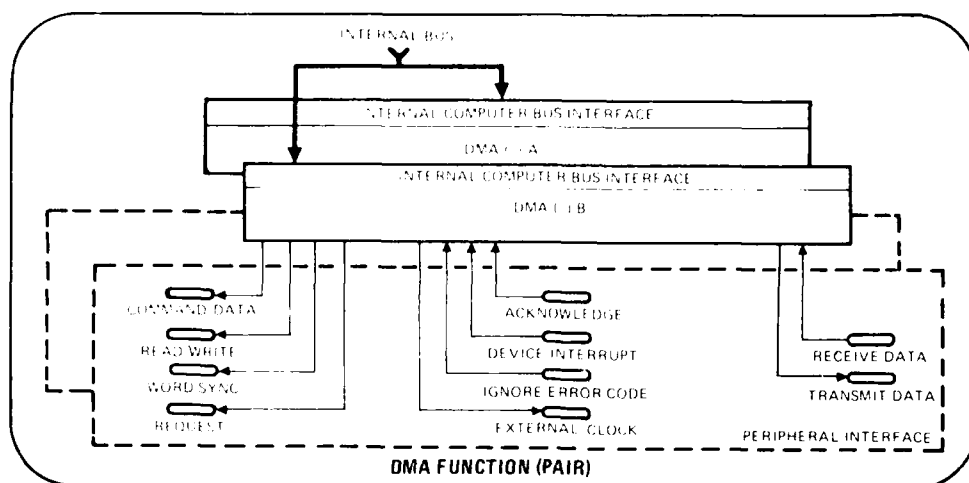
Use of error code is optional. If the ignore error code line is statically set low by the peripheral the error code must be appended to each word transferred over the DIU/Peripheral interface. In this case each word transferred will contain the 16 message bits of the serial bus word appended with eight code bits. The code transferred across the DIU/Peripheral interface assumes a 32 bit word containing zeros in bit positions 0-15 (most significant 16 bits) and the data contained on the data lines, in bit positions 16-31. The code bits are determined by a shortened cyclic code structure which is defined by the generator polynomial, $q(x) = (x^6 + x + 1)(x^2 + x + 1)$. When the line is in the high state, error code is not used during any transfers. For this case, the least significant eight data lines are not used.

As described, each DIU function consists of two powered DIU modules, one of which shall be on-line in the current configuration of the serial data bus. The user may employ one peripheral attached to both DIU modules with monitoring to determine which DIU module is on-line and/or the ability to accept a switch-over command from either DIU module. The user may also have two peripherals, one attached to each of the DIU modules or they may be cross-strapped. The DIU function will preserve error code bits, if provided by the user, or will append them, if not provided by the user, before it sends any information over the serial data bus. The error code will be the 8-bit code defined previously. The user has the responsibility of providing error detection for bad or missing output from the DIU modules. On detection of either of these conditions, the user can switch to the alternate DIU and produce a device interrupt. It can also cease to respond to further inputs. Should this occur, the FTSC can detect the state through software timers which can be provided in the specific device handling programs.

As previously described, the DIU pair is housed in the peripheral facilities which must accommodate both the dual-redundant serial bus link and the dual-24 bit parallel interfaces.

The DIU/Peripheral parallel interface signals are: low logic state or data value 0, $0V \pm 1.0V$; high logic state or data value 1, $10V \pm 1.0V$. The interfacing device shall present an impedance of at least 100K ohms and at most 50 pf. Its logic threshold shall be between 4V and 7V. It shall be able to drive into an FTSC input impedance of 10K ohms and 50 pf. Transitions between logic levels must be less than 200 nsec.

DIRECT MEMORY ACCESS (DMA)



Each Direct Memory Access (DMA) function provides an external user with a high-speed dual-redundant direct channel to the FTSC Main Memory. The DMA functions are program controlled and are capable of transferring blocks of data consisting of up to 4096 words from or to a specified Memory Module. Prior to a data transfer, the selected DMA is provided with a start-address, a stop-address, and an indication of data flow direction. The start-address is incremented each time a data word is transferred. When the incremented start-address equals the stop-address, the DMA ceases data transmission and raises an end-of-block interrupt to the CPU. User interrupts are signalled to the CPU via the DMA.

Data transfers between the DMA and its associated peripheral are effected using eight (8) Transmit Data lines, eight (8) Receive Data lines, six (6) handshake control lines and one (1) static control line. Four of the control lines emanate from the DMA: Request, Read/Write, Command/Data, Word Sync; and three from the peripheral: Device Interrupt, Acknowledge and Ignore Error Code.

The Request line is raised by the DMA to initiate a transfer to or from the peripheral. When this line is raised, the peripheral must act upon the current state of the other control lines.

The Read/Write line controls the direction of the transfer over the DMA/Peripheral interface. The line is set to the high state when the transfer is from the peripheral to the DMA and the line is set low when information is to be transmitted from the DMA to the peripheral.

The Command/Data line enables the peripheral to distinguish between command/status information and data. If this line is in the high state, information transferred from the DMA to the peripheral, and to the DMA from the peripheral, is defined as command and status, respectively. If the line is at the low state, information transferred in either direction is defined as data.

Information sent between a peripheral and the DMA is transmitted as 8-bit bytes. Five consecutive byte transfers are required to define a 40-bit data bus word. Word synchronization is accomplished by the DMA setting the Word Sync control line high during the transfer of the most significant byte (which is always transferred first).

The fifth byte transfer consists of the error code portion of the bus word, provided that the Ignore Error Code line is set low. If the peripheral user has selected not to append error code on data transfers across the DMA/Peripheral interface, the Ignore Error Code line will be set high and the fifth byte transfer shall not be exercised.

The cyclic code used is that defined by the generator polynomial, $g(x) = (x^6 + x + 1)(x^2 + x + 1)$. When the peripheral is providing error code, code bits will be appended to data words transferred to and from the user (the DMA passes the code with the words received from the internal FTSC bus to the user). If the user is not providing the code, the DMA generates code on the words as they are transferred into the FTSC across the interface.

When the DMA initiates a transfer, it raises the Request line. If the Read/Write line is set high, the peripheral places its data or status on the receive data lines and raises Acknowledge. If the Read/Write line is set low, the peripheral accepts the data or command present on the transmit data lines before it raises Acknowledge.

If the peripheral requires service, it raises the Device Interrupt Line. In response, the DMA sets the Command/Data and Read/Write line high and then initiates the byte transfer sequence. The DMA raises request prior to each byte transfer and waits for an acknowledge before proceeding. Word Sync is set high during the first byte transfer only.

The FTSC DMA is designed such that user interrupts are not serviced until all initiated word transfers have been completed. In the case of a transfer to/from the external data bus, each byte of the word must be transferred prior to servicing an interrupt.

Each of the two DMA functions consist of an active module and an inactive; standby module. The user may employ one peripheral attached to both DMA modules with monitoring to determine which DMA module is on-line and/or the ability to accept a switch-over command from either DMA module. The user may also have two peripherals, one attached to each of the DMA modules or they may be cross-strapped.

In addition to the control and data lines, each DMA provides a clock line for use by the external device. The clock frequency is the same as that of the internal FTSC clock.

Two peripheral interface options are available. When the FTSC and the peripheral are mechanically contiguous, standard log levels will be utilized at the interface: low logic state or data value 0, $0V \pm 1.0V$; high logic state or data value 1, $10V \pm 1.0V$. The interfacing device shall present an impedance of at least 100K ohms and at most 50 pf. Its logic threshold shall be between 4V and 7V. It shall be able to drive into an FTSC input impedance of 10K ohms and 50 pf. Transitions between logic levels shall be less than 200 nsec.

Where the peripheral-user is located remotely from the FTSC, a special interface is used which provides isolation.

RECOVERY CHARACTERISTICS

Fault recovery, the resumption of normal processing following the detection of a hardware failure, is accomplished by the complementary actions of dedicated hardware, a recovery program and program rollback.

The recovery program constitutes the "brain" of the recovery system. It carries out the fault diagnosis and reconfiguration strategies. The performance of the FTSC fault recovery system depends mainly on the recovery program: its features, its strategies, its implementation.

There are seven basic steps to recovery. They are:

1. Initial Fault Diagnosis
2. Identify Faulty Module
3. Determine Reconfiguration Strategy
4. Perform the Reconfiguration
5. Condition the new Element
6. Recover Elapsed Time
7. Rollback Application Programs

Steps 1, 6 and 7 are performed for all faults; the others are used as needed. The first step analyzes CCU information to determine the state of the machine, the type of fault and whether the fault occurred during recovery from a previous fault. Then CPU information is used to identify the module that had control of the buses when the fault was reported, and also the module that was being addressed. Device status information and testing, as necessary, are used to identify the faulty module. Available system resources are then examined. The most resource-efficient and subsequent-fault-resistant approach is taken. Once reconfiguration has been successfully completed; new elements, specifically memories and I/O devices, must be conditioned. Memory conditioning is based on the rollback strategy employed.

Briefly, this strategy calls for critical data to be doubly stored (duplex data), one copy in each of two software contiguous memory blocks. Conditioning consists of copying the duplex data from the partner software to the new memory. Status words of memories and I/O devices must be read to clear fault-indications and pending requests. When reconfiguration and conditioning have been completed, the CCU is notified, the hardened timer is stopped and "read". Once the system clock has been updated, the Executive is restarted. Using duplexed data, the Executive reactivates each previously active task at its last rollback address.

The recovery program performs the following basic functions:

- o RESOURCE MANAGEMENT - Maintains the System configuration tables. These indicate the current status of all software-selectable system components. A short term failure history is also kept.
- o COLD START - Identifies and processes a cold start. That is, establishes the initial system configuration. This is known as Anesthetic processing.
- o WARM START - Reestablishes the system configuration to the state existing prior to the power failure or circumvention shutdown. This is referred to as Alexic processing.

A DOUBLE GIMBALLED MOMENTUM WHEEL FOR PRECISION THREE-AXIS ATTITUDE CONTROL

by
Werner Auer
Head of the Engineering Department
for Gyroscopic and Space Products
TELDIX GmbH
6900 Heidelberg
West Germany

SUMMARY

For precision three-axis attitude control of space vehicles, a Double Gimballed Momentum Wheel (DGMW) as attitude actuator is a favorable approach. A high reliability DGMW of modular design with two momentum wheels and two direct drive ultra high resolution stepper motors plus pick-offs per gimbal axis (full redundancy; no caging during launch) was designed, built and qualification tested, together with an associated wheel and gimbal drive electronics. In addition, extensive system tests were performed on a three-axis air bearing table which proved the excellent control capabilities of this DGMW. Besides the measurement under stationary conditions with an accuracy of 18 sec of arc, also the transient behavior was studied. Due to the high torque capability (up to 0.2 Nm around the momentum wheel axis and up to 1 Nm around the gimbal axes) rapid and accurate pointing and repointing of spacecraft are possible. Therefore, the DGMW system can be designed into tracking loops to assure continuous pointing of spacecraft instruments with a rapid assignment capability.

1. INTRODUCTION

The success of 3-axis attitude control with fixed mounted momentum wheels in programs such as SYMPHONIE, INTELSAT V, ECS and others has proved the concept, where two axes (roll and yaw) normal to the spin axis of the momentum wheel are stabilized passively by gyroscopic action and where the control about the spin axis (pitch) is provided actively by momentum exchange (speed control of the wheel) with the satellite (Ref.1). For two-degree-of-freedom active control, V-configurations are successfully flown in the ECS and MARECS satellites. Single gimbal systems were studied also (Ref.2).

To obtain a three-degree-of-freedom active control several approaches are feasible. In this context, the sole utilization of thrusters (jets) is not taken into account because the necessary fuel supply cannot be established for longer missions. In contrast, momentum actuators need only electric energy which can be supplied by solar cells.

Two groups of approaches can be distinguished: Such employing one actuator, for instance, one reaction wheel per axis, and such utilizing only one momentum unit which is able to tilt the momentum vector in two axes and possesses in addition an alternation capability of the "length" of the momentum vector by changing the speed (like in the case of one-axis active control).

The latter approach promised a comparatively low mass, a high torque capability, a high reliability and excellent control properties, as has been found in many studies (Ref. 3, 4). Therefore, TELDIX has developed both, outer and inner gimbal DGMWs for ESTEC, respectively DFVLR. In the following, only the inner gimbal DGMW is treated which could be designed in an extraordinary compact way.

2. DESIGN OF THE DGMW

A modular design of the DGMW was chosen because of several reasons, of which the following are named:

- Adaptability concerning different requirements
- Availability of already flight proven momentum wheels, offering a wide range of angular momentum
- Qualification and life test results of gimbal drive motors and bearing arrangements
- Manufacturing, qualification and testing advantages, leading also to adequate costs

The photograph in Fig. 1 shows the DGMW, described in the following to a certain extent. It consists of an inner gimbal arrangement with the bearings, motors and pick-offs, and two wheels which may be operated simultaneously or one of them kept as a cold redundant unit. The interface to the wheels is standard; the interface to the satellite can be designed to customer's requirements.

The Wheel Drive Electronics (WDE) and the Gimbal Drive Electronics (GDE) are built as separate boxes and can be designed to defined interface requirements.

3. INNER GIMBAL ARRANGEMENT

Because the momentum wheels are designed without any caging or load relief mechanism (Ref.5), it was also intended to avoid any caging for the inner gimbal arrangement. This makes a symmetrical, rigid and stiff design approach mandatory. In the photograph (Fig. 2) the inner gimbal system is depicted which shows one of the two wheel interfaces, one of the two basic axle interfaces to the satellite in the foreground (which must be completed by a ring or a similar structure to solidly fix the two axles to each other, compare Fig. 1) and the four bearing, drive motor and pick-off assemblies, two for each axis.

Both gimbal axes are very rigidly defined by two preloaded pairs of ball bearings in a face-to-face arrangement. These bearings are operated in the vacuum found in satellite applications. To avoid any lubrication problems a concept employing ceramic balls instead of steel balls in the bearings without any lubricant was qualified. With these dissimilar materials no cold welding takes place. Also wear is negligible due to the small accumulated angular movement over mission time, as could be demonstrated in accelerated life tests, simulating a ten-years mission.

These findings are very important because the bearing suspension is the only critical element not provided in a redundant manner. Due to these test results and the environmental tests simulating launch loads, the bearings can be practically treated in terms of reliability like a part of the satellite's structure.

The motors are direct drive stepper motors, needing no extra gearing; they are free-wheeling if not energized. This is essential for the redundancy concept: No special mechanism is necessary for switching from the primary to the redundant unit.

The motors - having the same mechanical interface - are available in two versions, a tooth type and a friction type motor. The tooth type motor has a basic step angle of 2 min of arc which can be subdivided further by electronic means. The advantage is that - starting from a zero reference - by step counting the actual angle can be determined. The friction type motor operates with a step angle of about 10 sec of arc, depending on load.

Both motors were tested - also life tested - extensively in air and thermal vacuum. The wear in the gold plated contact zones between motor stator and motor rotor (teeth resp. friction surfaces) was very small and acceptable after excessive life tests.

Like other motors, this motor consists of a stator and a rotor. The stator of the friction type motor is designed as a ring with a circumferential conical surface (Fig.3) and is provided with equally spaced electromagnets.

The rotor needs no extra bearings but is suspended by the gimbal bearings and consists of a diaphragm which is equipped with a segmented soft iron ring (the edge of which can be seen in Fig. 2 at the extreme right). If a coil is energized, the stator attracts the rotor at the location of the coil, and stator and rotor are connected to each other by friction force (the coil with 180° offset is also energized thus doubling the torque capability). Now, if the coils are energized in a sequence, the contact zones between stator and rotor are moving around. Because of the small difference of the actual radii of stator and rotor (in other words: the small angle between the plan of the stator surface and the connection line rotor suspension - contact area of the rotor to the stator) a small advance (step) of the rotor occurs with each switching which is transferred to the axis by the diaphragm. Depending on the sense of the sequence, the gimbals can be driven in both directions. As already mentioned, if none of the coils is energized, the rotor is free to move. Thus the redundant motor on the other side of the gimbal axis can be operated.

The tooth type motor has a great number of teeth on the circumference of the stator and rotor with a difference in the number of one or more teeth. In this case, the stator has 720 teeth, the rotor 721. Therefore, after one "electrical revolution" the advance is one tooth. The stepwidth is the advance of one tooth divided by the number of coils.

It should be mentioned that these motors are well suited for digital control (Ref.6).

Presently, a microsyn pick-off is used which assures an accuracy of better than 2 min of arc over the whole gimbal rotational range of $\pm 7.5^\circ$. The resolution is adequate to the stepwidth.

The Gimbal Drive Electronics, not to be discussed here, "translates" the motor and pick-off formats into the formats required by the attitude control system.

4. MOMENTUM WHEEL

To constitute a DGMW, all IEDIX momentum wheels, presently available up to a nominal angular momentum of 70 Nms, can be mounted on the Inner Gimbal System. Due to the arrangement, the two wheels belonging to a DGMW must be operated relative to the wheel's interface in opposite rotational sense.

In Fig. 4 a sectional drawing of the momentum wheel is shown. It is a modular design with five subassemblies: Housing with mechanical interface, ball bearing unit, flywheel m.s., motor rotor and motor stator with commutation electronics.

The housing is a lightweight design with a center stud bolt to carry the loads exerted by atmospheric pressure after evacuation to about 1 mbar. The housing is black painted for thermal radiation.

The ball bearing unit is the most critical and significant subassembly of the momentum wheel. It is designed to meet the launch load requirements without any caging or load relief and to operate subsequently for missions of more than 10 years with high reliability (Ref.7). In the meantime, more than 35 years of accumulated flight experience (with one wheel operating more than 8 years) without any failure are a good proof for the reliability; in addition, on a total of nearly 50 years of life test results can be counted.

The flywheel mass is linked to the ball bearing unit by spokes. Resonance step-ups are limited by a special damper mechanism. The brushless and ironless DC motor is a low-loss high reliability design which also exhibits redundant features.

A Wheel Drive Electronics with torque and/or speed control modes forms the electrical interface required by the attitude control system.

5. MAIN DATA OF THE DGMW

In the following, the main data of the DGMW are listed. Concerning the mass they are based on a nominal angular momentum of 50 Nms for each of the two wheels. Tested were DGMW's with 20 Nms and 50 Nms wheels. As already mentioned, presently up to 70 Nms angular momentum wheels are available. The impact of wheels other than 50 Nms concerns mainly mass and power consumption.

Dimensions	
. Diameter	< 380 mm (envelope incl. tilting)
. Height	< 335 mm
Mass	
. Two wheels	22.4 kg
. Gimbal system	15.2 kg
. Gimbal system	7.2 kg
Wheel	
. Angular momentum	50 Nms
. Speed	4600 min ⁻¹ ± 10 %
. Reaction torque, max.	> 0.1 Nm
. Power	≤ 8 W
steady state	
at max. torque ..	≤ 80 W
Gimbal System	
. Tilting capability	7.5° each direction
. Torque capability	0.7 Nm
. Step angle	10 arc sec nominal, depending on torque
. Step frequency	≥ 130 Hz with 50 Nms wheels
. Pick-off range	± 8°
. Pick-off resolution	≤ 1 arc min
. Pick-off linearity	± 1 %
. Power	≤ 12 W
Environmental Capability, Qualification Level	
. Temperature	- 15 °C to 50 °C
. Vibration .. sinusoidal	10 to 2000 Hz 6 g
random	20 to 60 Hz 0.05 g ² /Hz
	60 to 300 Hz + 3 db/oct to 0.25 g ² /Hz
	300 to 1200 Hz 0.25 g ² /Hz
	1200 to 2000 Hz - 6 db/oct
. Linear acceleration	23 g
Reliability	0.998 (7 years)

6. TEST RESULTS

Together with an appropriate Wheel Drive Electronics (WDE) and a Gimbal Drive Electronics (GDE), a variety of tests at different levels were conducted. These tests comprised environmental and functional tests of the DGMW, including WDE and GDE - which are only shortly commented here - and system integration tests at MBB (Messerschmitt-Bölkow-Blohm) and DFVLR.

6.1 Functional and Environmental Tests

Successful functional and environmental tests were performed within the scope of a national television satellite program. It should be noted that the environmental tests were performed without any load relief or caging provisions. These tests verified the data specifications of which the main are accumulated in para 5.

6.2 System Tests

System tests were conducted first on an electrical integrator model, including a digital attitude and orbit control on-board computer. Additional, detailed and some results of these tests can be found in Ref.8. They showed an accuracy potential of better than 0.01 degree in pitch, yaw and roll due to the active control of three axes. Interesting is also the suggestion in Ref.8 to utilize radio frequency sensors in addition to earth infrared sensors not only for antenna pointing but for the attitude control of a spacecraft. As the infrared sensors use the Earth as a reference, as a reference, the direction to the centre of the Earth is defined since the beam direction of the antenna to a earth-stationed beacon is different from that, even a yaw signal can be generated. This seems to be very important because these tests proved that the achievable attitude accuracy depends mainly on the sensor accuracy.

Extensive 3-axis attitude control tests were prepared and performed on the 3-axis air bearing table of the DFVLR. The employed stepper motor gimbal actuation is well suited for direct digital control (Ref.9). The approach described there was implemented. The associated internal DFVLR report is listed under Ref.10 and should be available upon request. These tests were conducted partly with an optical reference to eliminate errors introduced by the presently available sensors and to investigate the accuracy potential of 3-axis active attitude control with a DGMW.

Results of the steady state operation are shown in Fig. 5. As can be seen, the achieved accuracy in roll and yaw is about 18 arc sec, roughly the double of the step angle of the friction type stepper motors. This error is a little higher than the expected error of 1 step and could be attributed to an error in the signal transfer between air bearing table and ground.

Attitude manoeuvres were performed, too. Input was a defined change of the attitude angle in one axis. The response to the other two axes (cross-coupling) was recorded also. In Fig. 6, the plots for a pitch and for a roll input are shown. The deviations are below 40 arc sec.

Last not least, simultaneous unloading cycles in roll and yaw can be recognized in Fig. 7. The input signals were reaction torques generated by gas jets on the air bearing table about the roll and yaw axes which are compensated by a tilting of the gimbals about the appropriate axis. The resulting attitude errors are again very small. Only during the change of slope, higher peaks are visible which are still relatively small (80 arc sec). These results compare favorably with those described in Ref.11.

Further tests on this air bearing table were aimed to studying the influence of flexible structures (e.g. solar panels) on attitude control accuracy. The satellites momentum is simulated by additional reaction wheels about each axis. These wheels could also be utilized to introduce disturbance torques (Ref.12).

7. POINTING, ACQUISITION AND TRACKING

In addition to attitude control with a high gyroscopic stiffness against disturbing torques, a DGMW can also fulfill at the same time the task of an attitude actuator for tracking loops of high accuracy to assure continuous pointing of instruments of a space vehicle to moving targets with the capability of rapid reassignment to other targets.

According to the case described in the first section of the previous para, error signals, for instance, generated by a control scan or a manoeuvre principle, can be fed into the attitude control loops. As has been demonstrated during the system tests, static and dynamic residual errors are smaller than some 10 sec of arc.

With the presently available torque capability of 0.7 Nm for a spacecraft in the 700 kg m² moment of inertia range, an angular acceleration of 1 mrad/sec² can be achieved. This means that after 10 s an angle of about 6 degrees can be accumulated.

If higher torque capabilities are required (high slew rate requirement) and/or high moment of inertia of a spacecraft or a section of a spacecraft, as those presently specified, gimbal drive motors with higher torque output must be provided.

The ultimate torque limiting factor is the tilting torque capability of the momentum wheel. With some derating the torque capability of its ball bearing unit is in the order of 200 Nm. Due to the light weight construction, the spokes of the flywheel mass are the limiting factor for the tilting torque. In the present design a 20 Nm capability can be taken into account, which is a factor of 30 compared with the present capability.

Around the wheel axis the torque capability is mainly limited by power dissipation. For the presently specified max. motor torque of 0.1 Nm a power of 80 W must be provided which is rising about proportionally with the motor torque. But it should be mentioned here that normally only in two axes a high torque is required.

8. CONCLUSION

For active attitude control of spacecraft several attitude actuator systems were taken into consideration (Ref.13). One of the preferred candidates for precision 3-axis attitude control is a DGMW.

A high reliability DGMW of modular design was developed and successfully qualification tested. System tests, also on a 3-axis air bearing table, proved the excellent accuracy potential of this DGMW for 3-axis active attitude control of spacecraft. In addition, DGMW systems can be designed into high accuracy tracking loops.

9. REFERENCES

- (1) U. Renner, "Attitude Control Requirements for Future Communication Satellites", VII IFAC Symposium on Automatic Control in Space, May 1976
- (2) J. Broquet, "Attitude Stabilization of Geostationary Satellites with a Single Degree of Freedom Angular Momentum Wheel System", IFAC Symposium on Automatic Control in Space, May 1976
- (3) M.G. Lyons, K.L. Lebsock, E.D. Scott, "Double Gimballed Reaction Wheel Attitude Control System for High Altitude Communication Satellites", AIAA Guidance, Control and Flight Mechanics Conference, 1971
- (4) H.L. Mork, "Synthesis and Design of a Gimballed Reaction Wheel Attitude Stabilization Package", AIAA Guidance, Control and Flight Mechanics Conference, 1971
- (5) W. Auer, "Design and Lubrication of Ball Bearing Units for Flywheels", European Space Tribology Symposium, Frascati, Italy, April 1975
- (6) E. Bruderle, Ch. Roche, H. Weingarten, "Attitude Control of Geostationary Telecommunication Satellite using Stepwise Motion of Gimbal Angles between Satellite and Flywheel", Symposium on Incremental Motion Control Systems and Devices, 1974
- (7) W. Auer, "Ball Bearings for Space Applications", Gyro Symposium, Stuttgart, 1981
- (8) E. Bittner, M. Wlaka, H.-D. Zago, "Attitude and Orbit Control System for Long-Life Communication Satellites", AIAA Guidance and Control Conference, Palo Alto, August 1978
- (9) G. Hirzinger, Th. Lange, "Direct Digital Attitude Control with a Double Gimballed Momentum Wheel, VII IFAC Symposium on Automatic Control in Space, May 1976
- (10) G. Hirzinger, Th. Lange, "Prüfung eines Stellrades (kardanisch gelagertes DRALLRAD) und einer Behelfs-elektronik auf einem Satellitenbewegungssimulator", Bericht A-552-77/10
- (11) J.J. Kalley, Jr. and H.L. Mork, "Development and Air Bearing Test of a Double Gimballed Momentum Wheel Attitude Control System", AIAA - 5th Communication Satellite Systems Conference, Los Angeles, April 1974
- (12) G. Heimfeld, G. Schulz, Th. Lange, "Physikalische Simulation von Satelliten mit flexiblen Auslagern", Deutscher Luft- und Raumfahrtkongress, Darmstadt, 1978, Vortrag-Nr. 78-120
- (13) W. Auer, "Ball Bearing Versus Magnetic Bearing Reaction and Momentum Wheels as Momentum Actuators", AIAA International Meeting and Technical Display, Baltimore, 1980, Paper No. AIAA-80-0911

10. ACKNOWLEDGEMENT

The author gratefully acknowledges the efforts of Mr. H.B. Schulz, Project Manager of the DGMW program.

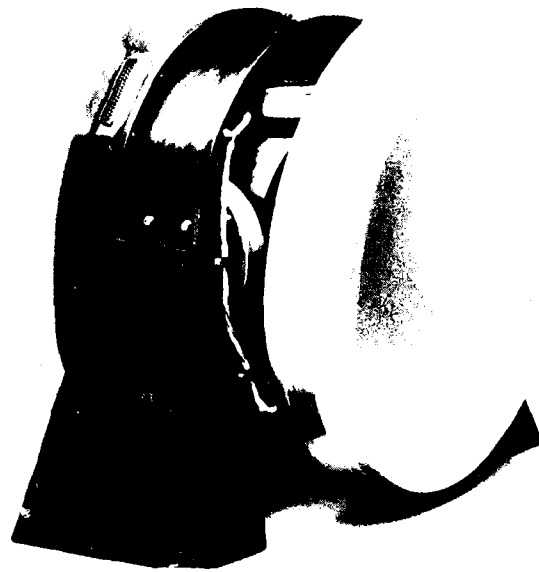


Fig. 1 Double Gimbaled Momentum Wheel

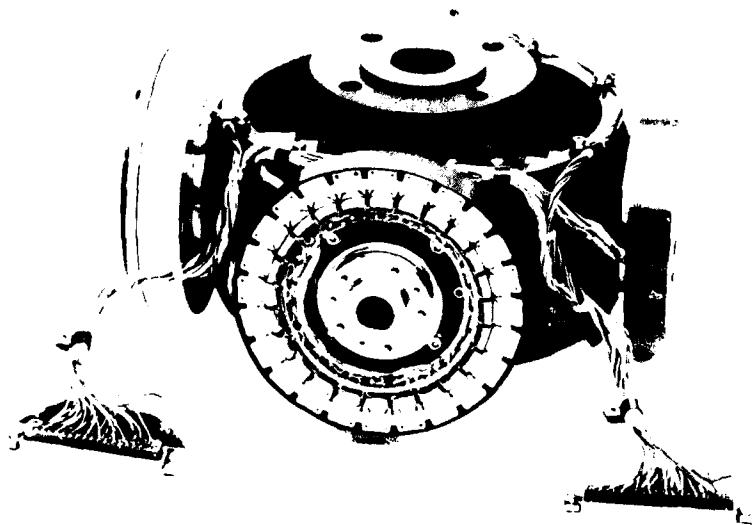


Fig. 2 Inner Gimbal Arrangement

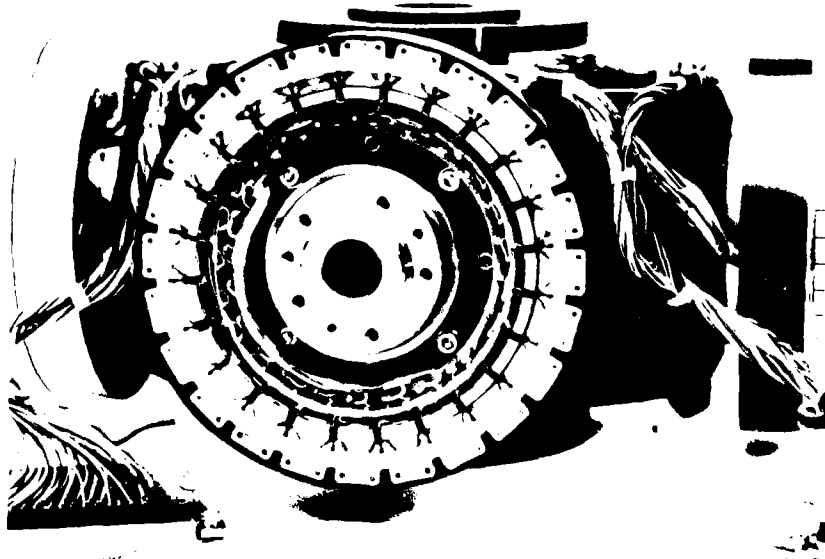


Fig. 3 Stepper Motor Stator

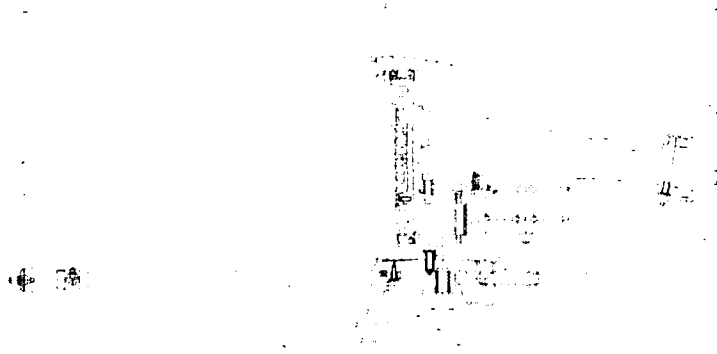


Fig. 4 Momentum Wheel

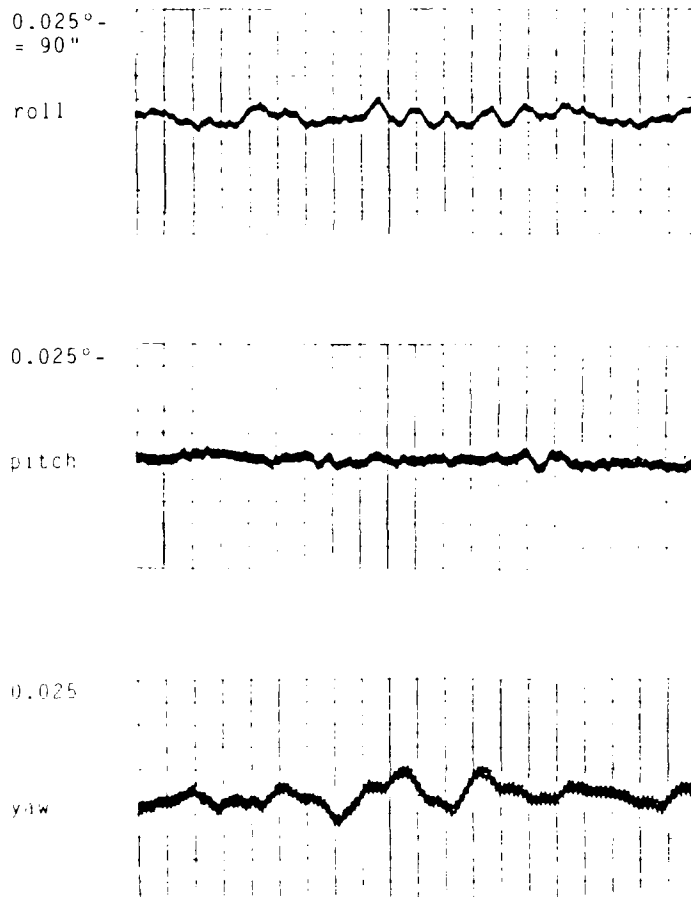


Fig. 5 Steady State Error

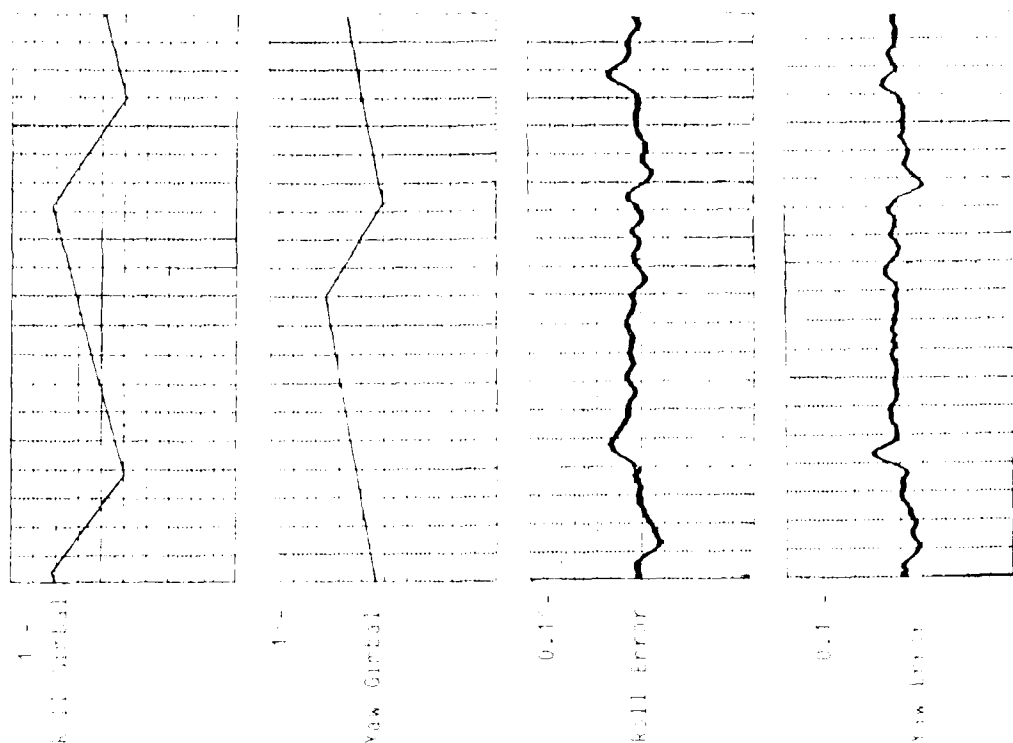
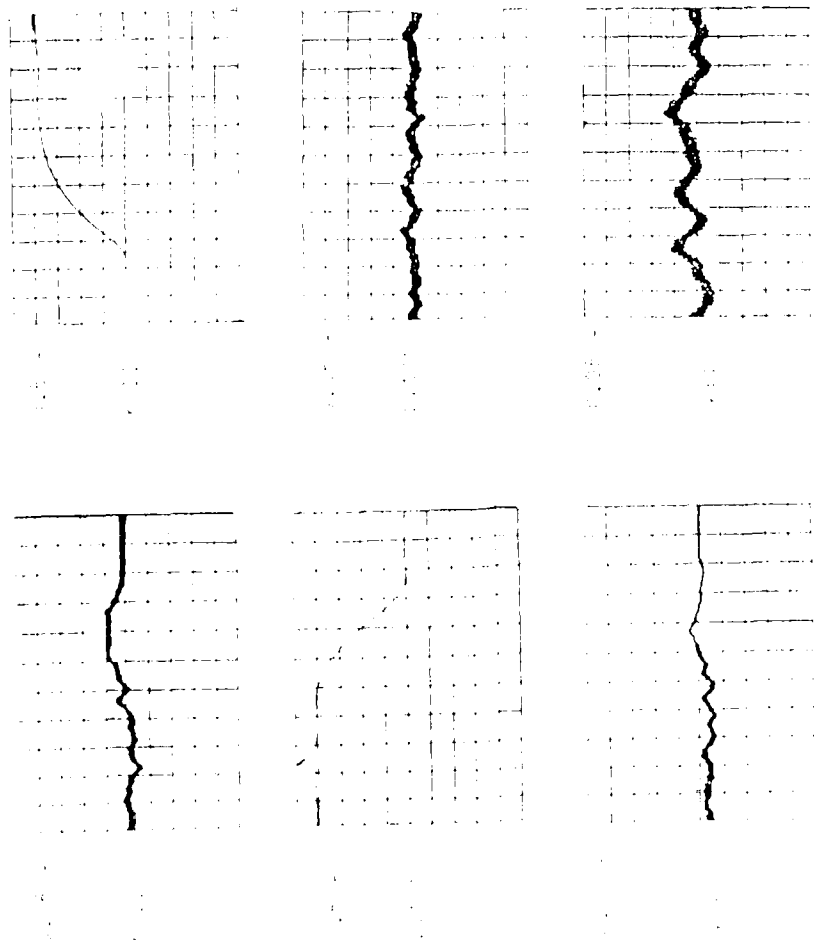


Figure 7 - Roll and Yaw Control Signals



CURRENT WORK ON HIGH-GRADE SPACE GYROS AT FERRANTI IN THE UK

by

John Wellburn
Navigation Systems Department
Ferranti PLC
Silverknowes, Edinburgh
Scotland

1. INTRODUCTION

Certain satellite missions require gyroscopes of very high accuracy for attitude measurement. Formerly, floated rate integrating gyros have been widely used. More recently in the USA, 2-axis DTG's have been used, and agencies are now beginning to consider what type of instrument to use in the future.

Whereas most US companies have tended to abandon floated gyros in favour of dry gyros (principally on grounds of manufacturing cost, we suspect) we at Ferranti have continued to develop and improve floated instruments. We are firmly of the opinion that for many satellite missions they will continue to offer the best performance.

Features such as fluid dissociation, which are rumoured to occur in floated gyros over a long period of time, can be controlled and minimised by care in design, assembly techniques, and cleanliness. In any event, the phenomena caused by fluid effects generally only affect g-sensitive performance aspects, which are of no relevance in orbital applications.

2. FLOATED GYRO ACTIVITIES

The instrument which has had the widest application in space in Europe is the Ferranti Type 125 Gyro. This has undergone continued development and improvement since its original license from Singer-Kearfott in the 1960's. It is a relatively small instrument, offering in-run drift stabilities in the range .001 to .00008 $^{\circ}$ /hr, depending on how it is used. It has been used in the IRAS, EXOSAT, and MIRANDA(X4) satellites, and in the EDDO, Black Arrow, and ARIANE launchers.

The question has arisen as to whether ball-bearing gyros are suitable for long-life satellite applications. Eight years ago, we started a life test of 12 of these gyros, under contract from RAE. The gyros were first subjected to an environmental test simulating a SCOUT launch. The gyros have been running continually in a partial vacuum environment, with performance checks and stop starts typically every 1-2 months. They have all now reached over 60,000 hours running without failure or degradation.

This performance justifies the claim that, for 5-year lifetime satellites at least, given suitable technology, and process control, ball-bearing gyros are perfectly adequate.

The life test is now continuing under the auspices of the E-SAT(OLYMPUS) programme.

The 125 Gyro has been extensively evaluated in terms of the spectral characteristics of drift, both long and short term. The in-run characteristics are a function of wheel speed. For many applications, the gyros are operated at low speeds around 117Hz, to maximise angular rate capability. Where this is not a limitation, they may be run at high speeds up to 350Hz, to give improved drift stability.

The short-term characteristics are exemplified by the data shown, obtained at NPL. The long-term characteristics are exemplified by the data obtained in the EADS programme, described by Max Norton.

To give an example of the characteristics at 350Hz wheel speed, the following figure shows the results of g-insensitive runs on both 125 ball-bearing and 127 gas-bearing gyros from the life test.

It can be seen that the g-insensitive runs for the two products are very similar. The main advantages are

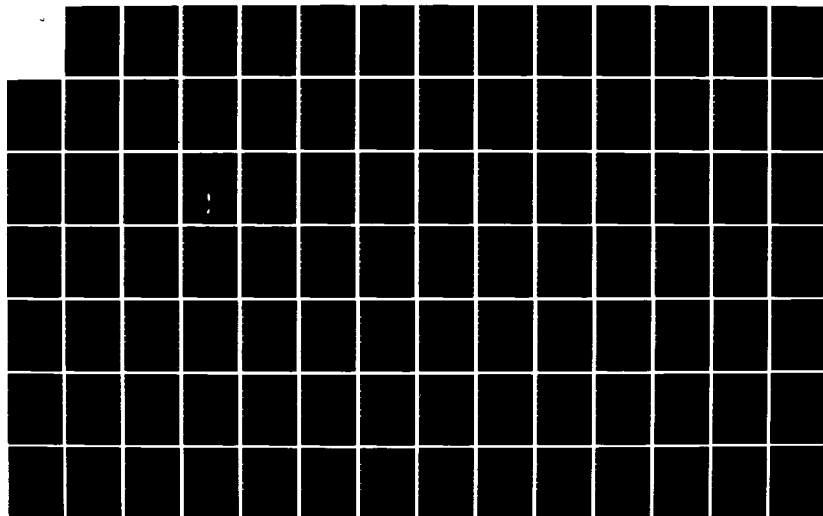
AD-A141 969

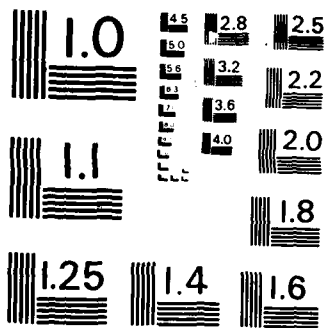
CONFERENCE PROCEEDINGS ON GUIDANCE AND CONTROL
TECHNIQUES FOR ADVANCED SP. (U) ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY... JAN 84
AGARD-CP-358 F/G 22/3

2/5

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

OVERCOMING UNOBSERVABILITY IN THREE-AXIS STABILIZATION OF SATELLITES

Dr. Lionel PASSERON
System Engineer
AEROSPATIALE
100 Bvd du Midi
B.P. 99
06322 CANNES-la-BOCCA Cedex
France

Dr. Claude BOZZO
Scientific Director
Compagnie de Signaux et
d'Entreprises Electriques
17 Place Etienne Pernet
75738 PARIS Cedex 15
France

Abstract : An optimal solution is provided to the problem of attitude estimation for a three-axis stabilized satellite devoid of yaw sensor and undergoing significant disturbing torques. Influence of such torques on expectable system performances is analysed via a geometric approach. Estimation of satellite attitude is demonstrated to be systematically impaired by an error resulting from the system's torque-induced unobservability. The said error is then explicitly assessed as a function of the disturbing torques. Furthermore, assuming a priori knowledge of the disturbing torque amplitudes, an optimal observability canonical form is derived. Optimal attitude estimation is finally provided, together with the expression of the relative, thus minimized, systematic error.

Introduction

Stabilizing satellites by means of the bias momentum concept is known as a simple, reliable and cost-saving solution, particularly well suited to geostationary missions. Its main advantage, doing away with yaw sensor and gyros, is nevertheless impaired by loss of accuracy -due to unobservability problems- when the spacecraft experiences significant disturbances.

This matter is now to be paid all the more attention as ever improved fine pointing accuracy is demanded, while the constantly growing size of satellites exposes them to increasing disturbing torques.

This paper shows that unobservability can be overcome to a wide extent by means of an original geometric approach. Use of optimal estimation and control algorithms, easily implementable on a microprocessor, provides -even on spacecraft undergoing severe disturbing torques- accuracy performances almost as good as when the zero-momentum concept is used.

In spite of the confined application of the method described here to standard-type three-axis stabilized geostationary satellites its easy extendability to any type of large flexible spacecraft should be emphasized.

The concluding achievement of our work, the baselines of which are featured hereinafter, is the demonstration by a simulation software that accuracy performances can be improved by a factor two with regard to most other, classical methods.

For greater simplicity, this paper is clear of all calculations and theoretical developments which may be found in the three Appendices hereto.

1. Satellite description

The geostationary, three-axis stabilized spacecraft considered here uses the pitch-momentum concept. Its main constituents are :

- a central body with some fixed, rigid appendages
- sun oriented solar panels
- a fixed momentum wheel with momentum vector aligned along the pitch axis.

This wheel provides gyroscopic rigidity and is operated as a reaction wheel to control the pitch axis

- one IR sensor providing continuous noisy measurements of the roll and pitch angles
- one sun sensor providing, once or twice per orbit, noisy measurements of the yaw angle
- thrusters to ensure unloading of the wheel as well as control of the roll and yaw axes.

The features here-above, it must be emphasized, are shared by quite a number of various, now orbiting communications satellites, civil (ITV, OTS, MARECS,...) and military (FLEETSATCOM, etc...) alike.

2. Satellite dynamics

Appendix I hereto develops, via a quasi-Lagrangian formulation, a general method to obtain the non-linear equations of motion for a satellite presenting any number of rotating appendages. Such equations are subsequently decoupled to enable analysis of the rotational motions independently of the translational ones.

Then the spacecraft rotational equations are linearized about the spacecraft prescribed behavior to achieve a state-space representation of the system and to apply the now classical theories of optimal control.

2.1 Spacecraft non-linear rotational equations

These may read, with the notations used in Appendix I hereto, as (see Appendix I, section 2.4) :

$$\begin{aligned} (J_E + \sum_i I_n^i) \ddot{\Omega} + \Omega (J_E + \sum_i I_n^i) \dot{\Omega} \\ + \sum_i \ddot{\theta}^i (W^i I_n^i - I_n^i \tilde{W}^i - I_n^i \tilde{W}^i) \Omega \\ + \sum_i \ddot{\theta}^i I_n^i W^i \\ + \sum_i (\ddot{\theta}^i)^2 \tilde{W}^i I_n^i W^i = \tilde{M}_{(ex)} R \end{aligned} \quad (1.1)$$

with

$$J_E = J + M \tilde{d}_e \tilde{d}_e + \sum_i m^i [\tilde{d}^i \tilde{d}^i - (\tilde{D}^i + \tilde{d}^i)(\tilde{D}^i + \tilde{d}^i)] \quad (1.2)$$

2.2 Spacecraft linearized rotational equations

Prescribed behavior defines itself by the coincidence of spacecraft reference frame \mathcal{R} with the instant local orbital frame \mathcal{R}_L (see Appendix III, section 1.2).

This stems from the necessity of having the antenna beam constantly pointed over some predetermined spot on the earth. These linearized equations are expressed as (see Appendix I, section 3.3).

$$\ddot{\theta} = - [{}^0\tilde{\Omega}_0 + X] \dot{\theta} - [X {}^0\tilde{\Omega}_0] \theta + (J_E + \sum_i I_n^i)^{-1} T_R + (J_E + \sum_i I_n^i)^{-1} U_R \quad (2.1)$$

with

$$X = (J_E + \sum_i I_n^i)^{-1} \{ {}^0\tilde{\Omega}_0 (J_E + \sum_i I_n^i) - (J_E + \sum_i I_n^i) {}^0\tilde{\Omega}_0 + \sum_i \dot{\theta}^i (\tilde{W}^i I_n^i - I_n^i \tilde{W}^i - I_n^i \dot{W}^i) \} \quad (2.2)$$

3. System state space representation

Since control of the pitch motion appears as quite simple, this paper focusses on the processing of the roll-yaw equations.

If we denote by T_Z and T_X the components of the satellite-affecting disturbing torques on yaw and roll axis respectively, these equations appear as (see Appendix III, section 1.4) :

$$\begin{cases} \ddot{\varphi} = (\omega_x)\dot{\varphi} + y\dot{\psi} - (\omega_y)\dot{\psi} + x\ddot{\psi} - \frac{y}{H} T_Z + \frac{x}{H} T_X \\ \ddot{\psi} = -(\omega_y)\dot{\varphi} - z\dot{\varphi} + (\omega_z)\dot{\psi} - y\ddot{\psi} + \frac{z}{H} T_Z - \frac{y}{H} T_X \end{cases} \quad (3)$$

with φ and ψ roll and yaw angle respectively.

3.1 The classical, reduced-order model

The disturbing torques, since small, random disturbances are assumed, are usually idealized as white noises. The resulting state-space representation, referred to as the reduced-order model hereinafter, then may read as :

$$\left\{ \begin{array}{l} \frac{d}{dt} \begin{bmatrix} \varphi \\ \dot{\varphi} \\ \psi \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \omega_x & y & -\omega_y & x \\ 0 & 0 & 0 & 1 \\ -\omega_y & -z & \omega_z & -y \end{bmatrix} \begin{bmatrix} \varphi \\ \dot{\varphi} \\ \psi \\ \dot{\psi} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\frac{y}{H} & \frac{x}{H} \\ 0 & 0 \\ \frac{z}{H} & -\frac{y}{H} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \\ y = [1 \quad 0 \quad 0 \quad 0] \begin{bmatrix} \varphi \\ \dot{\varphi} \\ \psi \\ \dot{\psi} \end{bmatrix} + W \end{array} \right. \quad (4)$$

Measurement y is provided by the IR sensor while W accounts for sensor noise.

The reduced-order model, despite its robustness, is well known for its poor efficiency when faced with significant disturbances.

3.2 The proposed, augmented-order model

A tentative overcoming of the drawback here-above consists in taking advantage of the quasi-deterministic nature of the disturbances.

The disturbing torques, witness the graphs in Fig.1 and 2 attached hereto, may be adequately modeled as truncated Fourier series, say :

$$T_Z = T_{Z0} + \sum_n T_{Zn}$$

$$T_X = T_{X0} + \sum_n T_{Xn}$$

with :

$$T_{Zn} = a_n^i \cos(n\omega t) + b_n^i \sin(n\omega t)$$

$$T_{Xn} = a_n^i \cos(n\omega t) + b_n^i \sin(n\omega t)$$

ω : Earth angular velocity.

Although this paper confines itself to first harmonic approximation, it should be emphasized that the results in Appendices II and III can straightforwardly apply to any number of torque harmonics.

Therefore, an augmented-order state vector :

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is defined, with vector

$$x_1 = [\varphi \quad \dot{\varphi} \quad \psi \quad \dot{\psi}]^T$$

reflecting satellite dynamics and vector

$$x_2 = [T_{Z_0} \quad T_{X_0} \quad T_{Z_1} \quad \dot{T}_{Z_1} \quad T_{X_1} \quad \dot{T}_{X_1}]^T$$

reflecting the disturbances.

By means of relation (3), the differential equations governing vector x here-above are derived leading to the following augmented-order state space representation :

$$\begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ B_{22} \end{bmatrix} v \\ y = [C \mid 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + w \end{cases} \quad (5)$$

with :

$$\begin{aligned} A_{11} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ \omega_x & y & -\omega_y & x \\ 0 & 0 & 0 & 1 \\ -\omega_y & -z & \omega_z & -y \end{bmatrix} \\ A_{12} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{y}{H} & \frac{x}{H} & -\frac{y}{H} & 0 & \frac{x}{H} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{z}{H} & -\frac{y}{H} & \frac{z}{H} & 0 & -\frac{y}{H} & 0 \end{bmatrix} \\ A_{22} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -\omega^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -\omega^2 & 0 \end{bmatrix} \\ B_{22} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ C &= [1 \quad 0 \quad 0 \quad 0] \end{aligned}$$

The above augmented-order model is the model under analysis in this paper.

4. The augmented-order model and associated unobservability problems

Let us start with a reminder on the Hautus Lemma which will provide us with some physical insight into the system's observability properties.

Lemma : Let that time-invariant linear system (S') whose state space representation reads as :

$$\begin{cases} \dot{X} = A' X \\ Y = C' X \end{cases}$$

The pair (C', A') is completely observable if and only if the composite matrix $[A'^T - \lambda I \mid C'^T]$ has full rank for all complex numbers λ .

If the matrix is not full rank for $\lambda = \lambda_i$, then the mode associated with the eigenvalue λ_i is not observable. The dimension of the unobservable subspace is equal to the rank deficiency.

Complete observability of the reduced-order model (4), fit for the no-disturbance case, can thus be readily demonstrated. On the contrary, the augmented-order model (5), fit for the case of significant disturbances, appears as partially unobservable.

As a matter of fact, matrix

$$A' = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

has three double eigenvalues which are $0, j\omega$ and $-j\omega$ whilst matrix

$$C' = [C \mid 0]$$

is a rank one matrix.

As a consequence of the Hautus Lemma, there is an unobservable subspace of dimension 1 for each eigenvalue $0, j\omega$ and $-j\omega$ mentioned here-above.

Physically speaking, this stems from the fact that roll and yaw disturbances are of the same frequencies, whilst only one measurement (the roll angle) is available.

5. Estimation of the satellite state-vector

The problem we are interested in is the construction of such an observer as provides, through use of the sole roll measurement, the best estimate (in a quadratic criterion sense) of the satellite's state vector x_1 .

The primary requisite for such observer construction is to derive a canonical form for observability of the said augmented-order model, having in mind that this model is endowed with the following properties:

- (H) a) (A_{11}, C) is completely observable
 b) matrices A_{11} and A_{22} have no common eigenvalue
 c) matrix A_{22} is diagonalizable

It should be noted that :

- verification of property (H.a) is straightforward since the reduced-order model has been proven completely observable
- verification of properties (H.b) and (H.c) are obvious.

For those systems which verify assumptions (H), Appendix II develops a geometric approach that does away with computation of the observability matrix.

The outcome is an optimal canonical form for observability which may read as :

$$\begin{cases} X_1 = x_1 + (N_1 + 2N_2) x_2 \\ X_{21} = S_1^T x_2 \\ {}^1X_{22} = {}^1S_2^T x_2 \\ {}^2X_{22} = {}^2S_2^T x_2 \end{cases} \quad (6)$$

with (see Appendix III, section 2.8) :

$$N_1 + 2N_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{H\omega} & 0 & \{D - \frac{(A^2 + B^2)X}{Z}\}\omega & 0 & \{\frac{CX}{Z} - 1\}A\omega & \{\frac{CX}{Z} - 1\}B \\ 0 & 0 & 0 & \{D - \frac{(A^2 + B^2)}{Z}\}\omega & \{1 - \frac{CX}{Z}\}B\omega^2 & \{\frac{CX}{Z} - 1\}A\omega \end{bmatrix}$$

$$S_1^T = [0 \quad \frac{1}{H\omega} \quad 0 \quad 0 \quad 0 \quad 0]$$

$${}^1S_2^T = \frac{1}{2HE} [0 \quad 0 \quad -A\omega \quad B \quad C\omega \quad 0]$$

$${}^2S_2^T = \frac{1}{2HE} [0 \quad 0 \quad B\omega \quad A \quad 0 \quad -C]$$

The above canonical form (6) is optimal to the extent that vector X_1 is that observable vector nearest (in a quadratic criterion sense) to the actual satellite state vector x_1 .

Furthermore, the assessment of the gap between the actual and approximate state vector, say $x_1 - X_1$, appears as an explicit function of the disturbing torques, say :

$$X_1 - x_1 = (N_1 + 2N_2) x_2$$

It is worthwhile noting that the unobservability problem is decoupled on each one of those characteristic subspaces of the system transition matrix that correspond to the different disturbing torque frequencies.

As evidenced in Appendix III (section 2.9), the state equations for the optimal observability canonical form appear as :

$$(7) \quad \begin{cases} \dot{X}_1 = A_{11} X_1 + \begin{bmatrix} 0 \\ \frac{x}{H} \\ 0 \\ -\frac{y}{H} \end{bmatrix} X_{21} + \frac{2E}{\omega^2 \{(A^2 + B^2)\gamma^2 + C^2 \delta^2\}} \begin{bmatrix} 0 & 0 \\ (yA\gamma^2 + \pi C \delta^2)\omega & -yB\gamma^2\omega \\ 0 & 0 \\ -(zA\gamma^2 + yC \delta^2)\omega & zB\gamma^2\omega \end{bmatrix} \begin{bmatrix} {}^1X_{22} \\ {}^2X_{22} \end{bmatrix} + (N_1 + 2N_2) B_{22} V \\ \dot{X}_{21} = S_1^T B_{22} V \\ {}^1\dot{X}_{22} = -\omega {}^2X_{22} + {}^1S_2^T B_{22} V \\ {}^2\dot{X}_{22} = \omega {}^1X_{22} + {}^2S_2^T B_{22} V \\ y = [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0] \begin{bmatrix} X_1 \\ X_{21} \\ {}^1X_{22} \\ {}^2X_{22} \end{bmatrix} \end{cases}$$

6. Optimal observer and quasi-optimal controller

These are built after prior sampling of the optimal observability canonical form (7) to enable implementation on micro-processors. The simulations use a 100 s sampling period and thrusters with a 4×10^{-3} Nms minimum impulse.

The observer is a plain Kalman filter.

The construction of the quasi-optimal controller is performed as explained hereinafter.

Let Φ be the transition matrix of the sampled system.

First, an asymptotic control gain is computed by minimization of a quadratic criterion.

Then, quantity

$$q_K = \|L\Phi\hat{x}_K\|$$

is assessed at each sampling step.

If q_K exceeds a user-selected threshold U (U , of course, must not be less than the minimum thruster impulse), control

$$U_K = -L\Phi\hat{x}_K$$

is applied to the spacecraft.

Now, if q_K is less than U , no control is applied on the satellite.

Thus, estimate \hat{x}_K is governed by an equation of the form :

$$\hat{x}_{K+1} = [I - KC] \Phi [\hat{x}_K + GU_K] + K y_{K+1}$$

with

$$U_K = -L \Phi \hat{x}_K \quad \text{if } q_K > U$$

$$U_K = 0 \quad \text{if } q_K < U$$

The simulations were performed with a threshold

$$U = 2 \times 10^{-2} \text{ Nms}$$

7. Simulation results

Figures 3 and 4, attached hereto, display the observer's performances when the actual disturbing torques acting on the spacecraft contain no harmonics greater than one. The actual yaw error, it must be stressed, is centered around the systematic error (which does not exceed 0.12°). Figures 5 and 6 display the behavior of the controlled spacecraft when undergoing actual disturbing torques together with the thruster commands.

Roll and yaw errors are found not to exceed 0.07° and 0.21° respectively. It is noteworthy that those currently flying satellites which are of the type analyzed here present a yaw error about 0.5° .

The thruster consumption values are quite reasonable, say 0.5 Kg annual for the roll thruster as against 0.14 Kg for yaw.

The annual number of thruster pulses, say 35,000 for roll, 11,000 for yaw thruster is quite consistent with the usually accepted reliability data.

Conclusion

Through a geometric approach of observability, this paper has derived :

- an explicit expression of the systematic estimation error induced by the disturbing torques
- an optimal canonical form for observability enabling the construction of an observer which gives the best possible estimation of the spacecraft's attitude
- an assessment of those optimal performances expectable from the fine pointing mode three-axis stabilization
- a highly efficient quasi-optimal controller,

all of the above for a satellite devoid of yaw sensor and stabilized by means of the bias momentum concept.

It is hoped that the method developed throughout this paper will meet practical applications since it appears as quite simple and easily implementable on microprocessors.

• References •

1. Alengrin, G., L. Passeron, (1982), *Minimizing Unobservability for Linear Systems with Unmeasurable Periodic Disturbances*, 6th IFAC Symp. on Identification and System Parameter Estimation, Washington D.C., USA
2. Bryson, A.E., Jr. and W. Kortum, *Estimation of the Local Attitude of Orbiting Spacecraft*
3. Dougherty, H.J., E.D. Scott, J.J. Rodden, (1968), *Analysis and Design of Whecon - An Attitude Control Concept*, AIAA, 2nd Comm. Satellite Systems Conf., San Francisco, Cal., USA
4. Hautus, M.L.J., (1969), *Controllability and observability conditions of linear autonomous systems*, Proc. Ned. Akad. Van Wetensch., serie A, vol. 72, n°5, 443-448
5. Ivens, R.P., A.W. Fleming, V.A. Spector, (1974), *Precision Attitude Control with a Single Body - Fixed Momentum Wheel*, AIAA Mechanics and Control of Flight Conf., Anaheim, Cal., USA
6. Morari, M., G. Stephanopoulos, (1980), *Minimizing unobservability in inferential control schemes*, Int. J. of Control, vol. 31, n°2, 367-377
7. Passeron, L., (1981), *Application du filtrage de Kalman au contrôle d'attitude de satellite*, Univ. de Nice
8. Passeron, L., (1983), *Dynamic Modeling of Flexible Spacecraft - A General Program for Simulation & Control*, 34th IAF Congress, Budapest, Hungary
9. Wonham, W.M., (1974), *Linear Multivariable control: A geometric approach*, Lec. Notes in Economic & Mathematical systems, Berlin, Springer 1974.

• ACKNOWLEDGMENTS •

For the interesting discussions we had on this work, we are indebted to J.J. DECHEZELLES, A. MORTELETTE, both of Aérospatiale-Cannes.

From G. ALENGRIN, Université de Nice and J.P. MONNIER, Aérospatiale-Cannes, we received invaluable technical support throughout our effort.

Finally, a special thought goes to Madame D. ARNUS, Université de Nice, and to P. PIERSON, Aérospatiale-Cannes, without whom this paper could never have been produced.

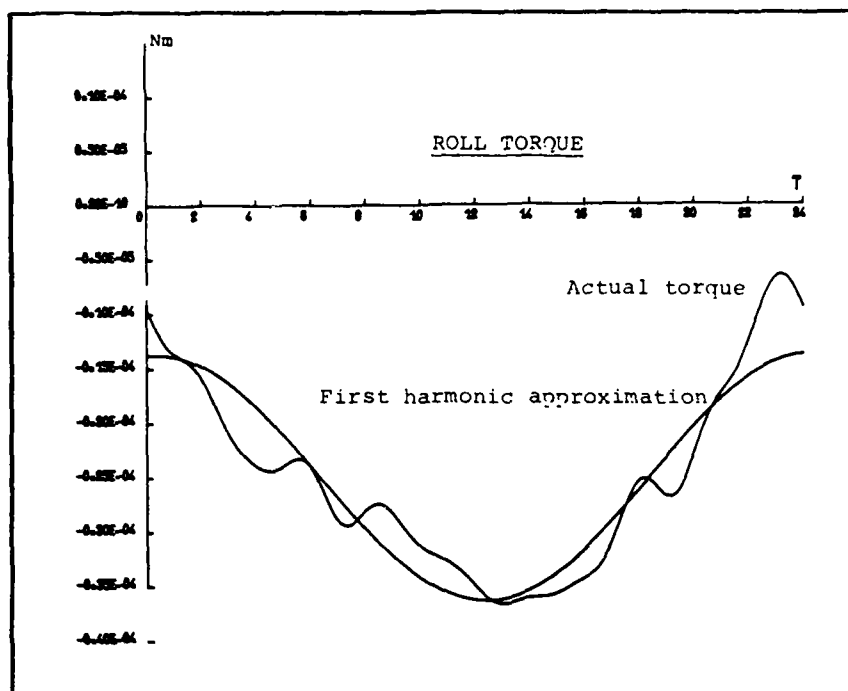


Figure 1

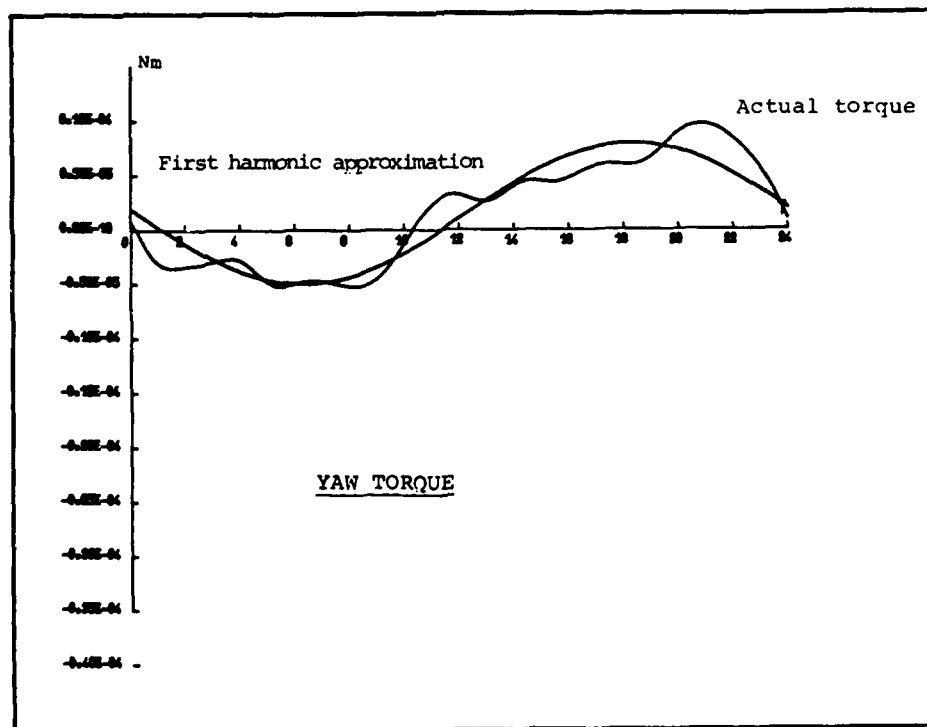


Figure 2

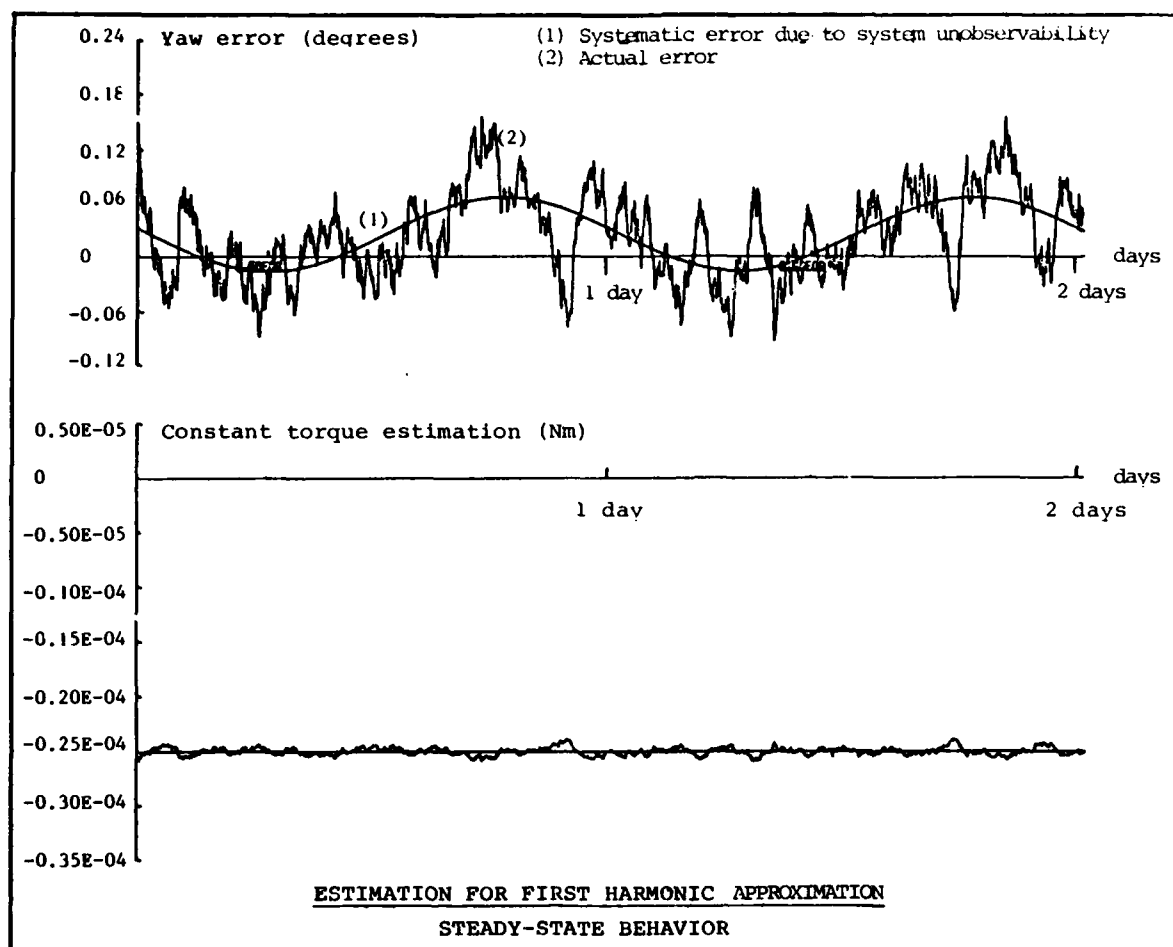


Figure 3

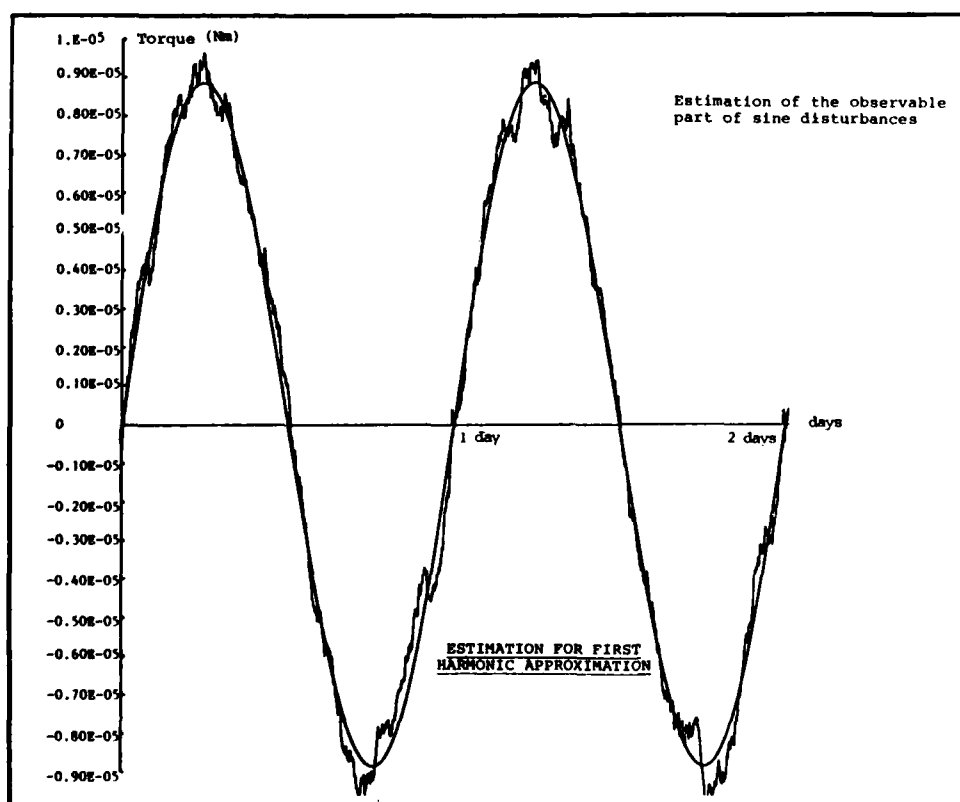


Figure 4

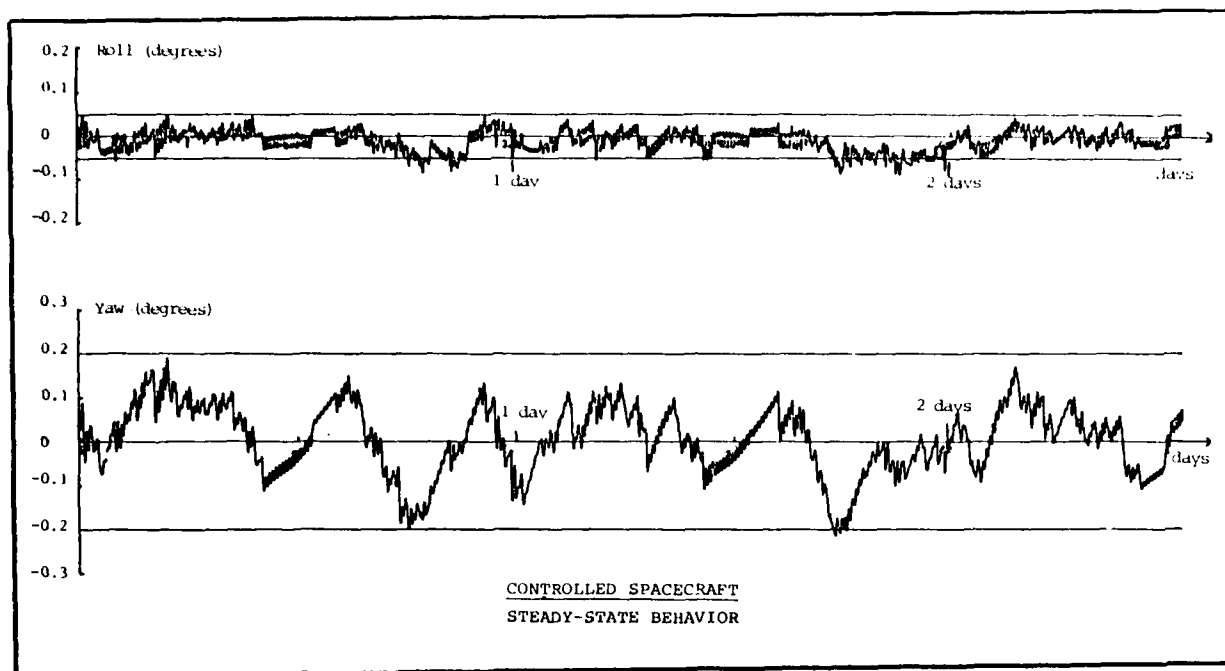


Figure 5

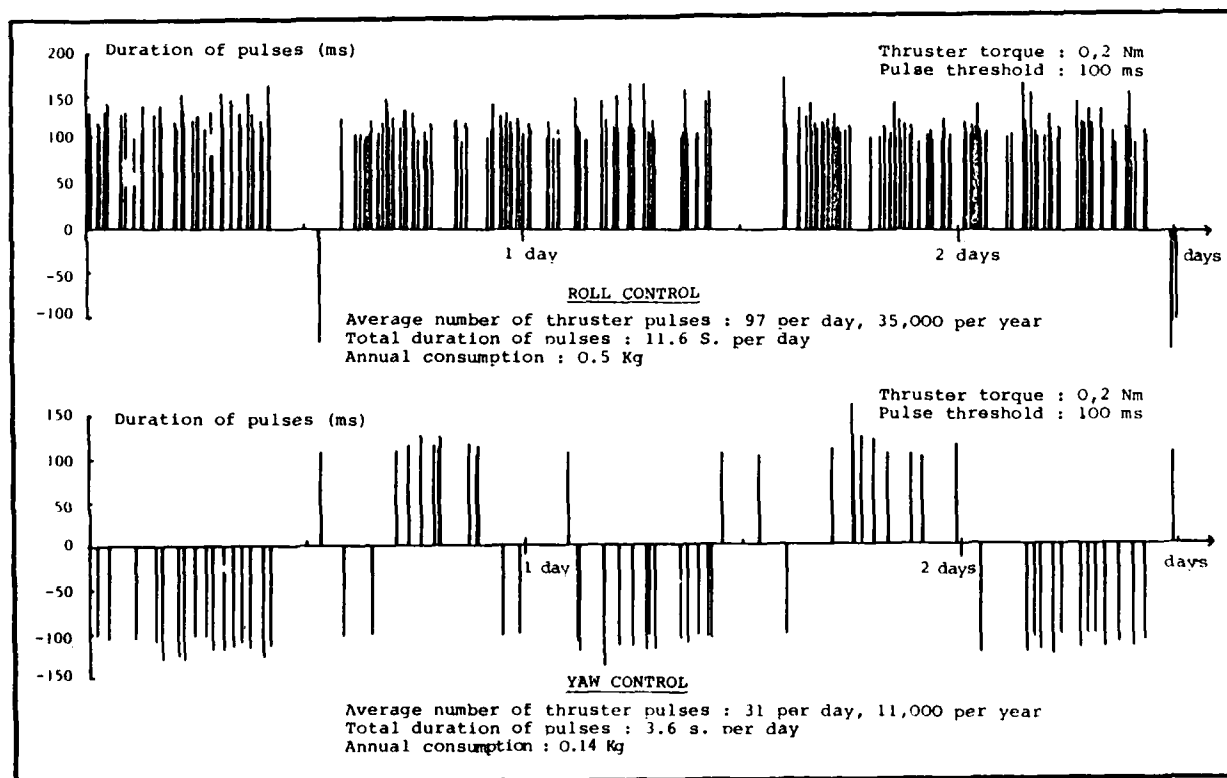


Figure 6

APPENDIX I.

DYNAMIC EQUATIONS FOR A SATELLITE PRESENTING SPINNING APPENDAGES
• Linearization around some prescribed behavior •

The non-linear equations of motion for a satellite presenting rotating appendages are first established through a Lagrangian approach.

A simple combination of the said equations, then, enables decoupling of the rotational from the translational motions.

Next, linearization around some prescribed behavior is carried out

It is assumed hereinafter, for greater simplicity, that the centers of gravity of all appendages lie on their rotational axes.

1. Spacecraft description

The type of spacecraft under consideration is comprised of $(n+1)$ parts, say :

- one main body S
- n rotating bodies (momentum wheels, solar arrays,...) S^i , $1 \leq i \leq n$

Let index i ($1 \leq i \leq n$) refer to all that is relative to body S^i (whilst body S has no index).

Then the following notations are adopted :

- $\mathcal{R} = (R, \vec{U}, \vec{V}, \vec{W})$ as the reference frame of body S
- $\mathcal{R}^i = (R^i, \vec{U}^i, \vec{V}^i, \vec{W}^i)$ as the reference frame of body S^i
- G (resp. G^i) as the center of gravity of body S (resp. S^i)
- m (resp. m^i) as the mass of body S (resp. S^i)
- M as the total mass of the spacecraft
- \vec{d} (resp. \vec{d}^i) as the vector \vec{RG} (resp. $\vec{R}^i G^i$)
- \vec{D}^i as the vector \vec{RR}^i
- J (resp. J^i) as the inertia tensor of body S (resp. S^i) relative to frame \mathcal{R} (resp. \mathcal{R}^i)
- I_n^i as the inertia tensor of body S^i relative to frame \mathcal{R}
- \vec{V}_R (resp. \vec{V}_{R^i}) as the absolute velocity of point R (resp. R^i)
- $\vec{\Omega}$ as the absolute angular velocity of body S
- $\vec{\Omega}^i$ as the angular velocity of body S^i relative to body S
- $\vec{F}_{(ex)}$ as the resultant external force acting on the spacecraft
- $\vec{M}_{(ex)R}$ as the resultant external torque (with respect to point R) acting on the spacecraft.

Conventionally frame \mathcal{R}^i is so defined that

- point R^i lies on the rotation axis of body S^i
- vector \vec{W}^i is co-linear with the same axis

Hence : $\vec{\Omega}^i = \dot{\theta}^i \vec{W}^i$
where θ^i is the angle of rotation of body S^i about own axis.

2. Spacecraft dynamic equations

The following three conventions are used hereinafter :

- P^i as the transition matrix from frame \mathcal{R} to frame \mathcal{R}^i
- X (resp. iX) as the components of vector \vec{X} referenced to frame \mathcal{R} (resp. \mathcal{R}^i)
- \tilde{X} as the order 2 skew symmetric tensor :

$$\tilde{X} = \begin{bmatrix} 0 & -(X)_3 & (X)_2 \\ (X)_3 & 0 & -(X)_1 \\ -(X)_2 & (X)_1 & 0 \end{bmatrix}$$

where : $X = [(X)_1 \quad (X)_2 \quad (X)_3]^T$ is any vector of \mathbb{R}^3 .

2.1 Formulation of the Lagrangian equations

Let T and T^i be the kinetic energies of bodies S and S^i respectively, and :

$$\bar{T} = T + \sum_i T^i$$

be the total spacecraft kinetic energy.

Let all vectors be expressed by their components relative to moving frame \mathcal{R} .

Let T and T^i , and thereby \bar{T} , finally be computed as functions of the said components. Then it can be shown that, relative to the moving frame \mathcal{R} , the Lagrange equations for the spacecraft may read as follows :

α) overall translational equations

$$\frac{d}{dt} \left[\frac{\partial \bar{T}}{\partial \vec{V}_R} \right] + \vec{\Omega} \frac{\partial \bar{T}}{\partial \vec{V}_R} = \vec{F}_{(ex)}$$

β) overall rotational equations

$$\frac{d}{dt} \left[\frac{\partial \bar{T}}{\partial \vec{\Omega}} \right] + \vec{\Omega} \frac{\partial \bar{T}}{\partial \vec{\Omega}} + \vec{V}_R \frac{\partial \bar{T}}{\partial \vec{V}_R} = \vec{M}_{(ex)R}$$

γ) rotational equation for spinning appendage S^i

$$\frac{d}{dt} \left[\frac{\partial T}{\partial \dot{\theta}^i} \right] - \frac{\partial T}{\partial \theta^i} = \mu_{(ex)}^i$$

with $\mu_{(ex)}^i$ as the resultant external torque (with respect to rotation axis of body S^i) acting on S^i .

2.2 Computation of the kinetic energies and of some useful quantities :

2.2.1 Expressions of kinetic energies T and T^i

Kinetic energy T may read as :

$$T = \frac{1}{2} V_R^T m V_R + \frac{1}{2} \Omega^T J \Omega + m \Omega^T \tilde{d} V_R \quad (1)$$

whilst kinetic energy of body S^i appears as :

$$T^{i*} = \frac{1}{2} V_{Ri}^T m^i V_{Ri} + \frac{1}{2} \Omega_A^{iT} J^i \Omega_A^i + m^i \Omega_A^{iT} \tilde{d}^i V_{Ri} \quad (2)$$

where Ω_A^i is the absolute angular velocity of body S^i .

When the vectors \tilde{V}_{Ri} , \tilde{d}^i and the pseudo-vector $\tilde{\Omega}_A^i$ are expressed by their components relative to frame \mathcal{N}^i , relation (2) here-above now becomes :

$$T^{i*} = \frac{1}{2} V_{Ri}^T m^i V_{Ri} + \frac{1}{2} \Omega_A^{iT} I_n^i \Omega_A^i + m^i \Omega_A^{iT} \tilde{d}^i V_{Ri} \quad (3)$$

where inertia tensor I_n^i of body S^i relative to frame \mathcal{N} is defined as follows :

$$I_n^i = p^i J^i p^{iT}$$

Combined use of equalities :

$$V_{Ri} = V_R - \tilde{D}^i \Omega$$

$$\Omega_A^i = \Omega + \tilde{\Omega}^i$$

and of relation (3) here-above would enable kinetic energy of body S^i to be expressed as a function of V_R , Ω and Ω^i , say :

$$T^i(V_R, \Omega) = T^{i*}(V_{Ri}(V_R, \Omega), \Omega_A^i(\Omega, \Omega^i)) \quad (4)$$

However, as will clearly appear later, T^i need not be explicitly developed.

2.2.2 Computation of some useful quantities

The quantities :

$$\frac{\partial T}{\partial V_R}, \frac{\partial T^i}{\partial V_R}, \frac{\partial T}{\partial \Omega}, \frac{\partial T^i}{\partial \Omega}, \frac{\partial T^i}{\partial \theta^i}$$

and :

$$\frac{d}{dt} \left[\frac{\partial T}{\partial V_R} \right], \frac{d}{dt} \left[\frac{\partial T^i}{\partial V_R} \right], \frac{d}{dt} \left[\frac{\partial T}{\partial \Omega} \right], \frac{d}{dt} \left[\frac{\partial T^i}{\partial \Omega} \right], \frac{d}{dt} \left[\frac{\partial T^i}{\partial \theta^i} \right]$$

which further enable assessment of the lefthandside terms of the Lagrange equations are now computed.

Such computation will utilize the identities :

$$\frac{\partial T^i}{\partial V_R} = \frac{\partial T^{i*}}{\partial \tilde{V}_{Ri}}$$

$$\frac{\partial T^i}{\partial \Omega} = \frac{\partial T^{i*}}{\partial \Omega_A^i} + \tilde{D}^i \frac{\partial T^{i*}}{\partial V_{Ri}}$$

which are readily derived from relation (4).

One obtains :

$$\frac{\partial T}{\partial V_R} = m V_R - m \tilde{d} \Omega$$

$$\frac{\partial T}{\partial \Omega} = J \Omega + m \tilde{d} V_R$$

$$\frac{d}{dt} \left[\frac{\partial T}{\partial V_R} \right] = m \dot{V}_R - m \tilde{d} \dot{\Omega}$$

$$\frac{d}{dt} \left[\frac{\partial T}{\partial \Omega} \right] = J \dot{\Omega} + m \tilde{d} \dot{V}_R$$

and :

$$\frac{\partial T^i}{\partial V_{Ri}} = m^i (V_{Ri} - \tilde{d}^i \Omega_A^i)$$

$$= m^i [V_{Ri} - (\tilde{D}^i + \tilde{d}^i) \Omega]$$

$$\begin{aligned} \frac{\partial T^i}{\partial \Omega} &= [I_n^i - m^i \tilde{D}^i \tilde{d}^i] \Omega_A^i + m^i (\tilde{D}^i + \tilde{d}^i) V_{Ri} \\ &= m^i (\tilde{D}^i + \tilde{d}^i) V_R + [I_n^i - m^i (\tilde{D}^i \tilde{d}^i + \tilde{d}^i \tilde{D}^i + \tilde{D}^i \tilde{D}^i)] \Omega + \theta^i I_n^i w^i \end{aligned}$$

$$\begin{aligned} \frac{\partial T^i}{\partial \theta^i} &= \frac{1}{2} \Omega_A^{iT} (\tilde{w}^i I_n^i - I_n^i \tilde{w}^i) \Omega_A^i \\ &= -w^{iT} \tilde{\Omega} I_n^i (\Omega + \theta^i w^i) \end{aligned}$$

$$\frac{\partial T^i}{\partial \dot{\theta}^i} = w^{iT} I_n^i (\Omega + \theta^i w^i)$$

$$\frac{d}{dt} \left[\frac{\partial T^i}{\partial V_{Ri}} \right] = m^i [\dot{V}_{Ri} - (\tilde{D}^i + \tilde{d}^i) \dot{\Omega}]$$

$$\begin{aligned} \frac{d}{dt} \left[\frac{\partial T^i}{\partial \dot{\Omega}^i} \right] &= m^i (\tilde{D}^i + \tilde{d}^i) \dot{V}_R \\ &+ [In^i - m^i (\tilde{D}^i \tilde{d}^i + \tilde{d}^i \tilde{D}^i + \tilde{D}^i \tilde{D}^i)] \dot{\Omega} \\ &+ \dot{\theta}^i (\tilde{W}^i In^i - In^i \tilde{W}^i) \Omega \\ &+ \dot{\theta}^i In^i W^i \\ &+ (\dot{\theta}^i)^2 \tilde{W}^i In^i W^i \\ \frac{d}{dt} \left[\frac{\partial T^i}{\partial \dot{\theta}^i} \right] &= W^{iT} \{ In^i (\dot{\Omega} + \dot{\theta}^i W^i) - \dot{\theta}^i In^i \tilde{W}^i \Omega \} \end{aligned}$$

2.3 Spacecraft dynamic model

The results given in sections 2.1 and 2.2 here-above enable the satellite's motion equations to be derived.

Setting :

$$d_e = \frac{m d + \sum_i m^i (\tilde{D}^i + \tilde{d}^i)}{M}$$

and after some manipulations, the said equations appear in the following forms :

α) overall translational equations

$$-M \{ \tilde{d}_e \dot{\Omega} + \tilde{\Omega} \tilde{d}_e \Omega \} + M \{ \dot{V}_R + \tilde{\Omega} V_R \} = \bar{F}_{(ex)} \quad (5)$$

β) overall rotational equations

$$\begin{aligned} &\{ J + \sum_i [In^i + m^i \tilde{d}^i \tilde{d}^i - m^i (\tilde{D}^i + \tilde{d}^i)(\tilde{D}^i + \tilde{d}^i)] \} \dot{\Omega} \\ &+ \tilde{\Omega} \{ J + \sum_i [In^i + m^i \tilde{d}^i \tilde{d}^i - m^i (\tilde{D}^i + \tilde{d}^i)(\tilde{D}^i + \tilde{d}^i)] \} \Omega \\ &+ \sum_i \dot{\theta}^i \{ \tilde{W}^i In^i - In^i \tilde{W}^i - In^i W^i \} \Omega \\ &+ \sum_i \dot{\theta}^i In^i W^i \\ &+ \sum_i (\dot{\theta}^i)^2 \tilde{W}^i In^i W^i \\ &+ M \tilde{d}_e \{ \dot{V}_R + \tilde{\Omega} V_R \} = \bar{M}_{(ex)R} \end{aligned} \quad (6)$$

γ) rotational equation for spinning appendage S^i

$$W^{iT} \{ In^i (\dot{\Omega} + \dot{\theta}^i W^i) + \tilde{\Omega} In^i \Omega + \dot{\theta}^i (\tilde{\Omega} In^i + In^i \tilde{\Omega}) W^i \} = \bar{M}_{(ex)}^i \quad (7)$$

2.4 Decoupling rotations from translations

Lefthandside multiplication of relation (5) by $(-\tilde{d}_e)$, followed by addition to relation (6), and then by setting

$$J_E = J + M \tilde{d}_e \tilde{d}_e + \sum_i m^i [\tilde{d}^i \tilde{d}^i - (\tilde{D}^i + \tilde{d}^i)(\tilde{D}^i + \tilde{d}^i)] \quad (8)$$

yield the decoupled overall rotational equations :

$$\begin{aligned} &(J_E + \sum_i In^i) \dot{\Omega} + \tilde{\Omega} (J_E + \sum_i In^i) \Omega \\ &+ \sum_i \dot{\theta}^i (\tilde{W}^i In^i - In^i \tilde{W}^i - In^i W^i) \Omega \\ &+ \sum_i \dot{\theta}^i In^i W^i \\ &+ \sum_i (\dot{\theta}^i)^2 \tilde{W}^i In^i W^i = \bar{M}_{(ex)R} \end{aligned} \quad (9)$$

3. Linearization of spacecraft rotational equations

Such linearization is performed around the prescribed spacecraft behavior.

3.1 Notations

All that refers to prescribed spacecraft behavior is denoted by a bottom right index o. Thus :

- $\mathcal{N}_o = (R_o, \vec{U}_o, \vec{V}_o, \vec{W}_o)$ designates frame \mathcal{N} for prescribed behavior.
- $\vec{\Omega}_o$ is the absolute angular velocity of frame \mathcal{N}_o .

Moreover, if \vec{X} is any vector, ${}^o\vec{X}$ denotes the components of vector \vec{X} relative to frame \mathcal{N}_o .

3.2 Linearization of absolute angular velocity

Suppose that frame \mathcal{A} remains close to \mathcal{A}_0 .

Let $\underline{\theta} = [\varphi, \theta, \psi]^T$

be the three pseudo Euler angles which enable characterization of frame \mathcal{A} relative to frame \mathcal{A}_0 .

Let P_0 be the transition matrix from frame \mathcal{A}_0 to frame \mathcal{A} . P_0 can straightforwardly be expressed as :

$$P_0 = I + \underline{\tilde{\theta}}$$

Then :

$$P_0^{-1} = I - \underline{\tilde{\theta}}$$

and for any vector \vec{X} :

$$\underline{X} = (I - \underline{\tilde{\theta}})^0 \underline{X}$$

which finally leads to :

$$\underline{X} = {}^0\underline{X} + {}^0\underline{\tilde{X}} \underline{\theta} \quad (10)$$

On the other hand, absolute angular velocity $\underline{\Omega}$ of frame \mathcal{A} equals the summation of the absolute angular velocity of frame \mathcal{A}_0 , say $\underline{\Omega}_0$, and the angular velocity of frame \mathcal{A} relative to frame \mathcal{A}_0 , say $\underline{\Omega}'$:

$$\underline{\Omega} = \underline{\Omega}_0 + \underline{\Omega}' \quad (11)$$

Since (cf. (10))

$$\underline{\Omega}_0 = {}^0\underline{\Omega}_0 + {}^0\underline{\tilde{\Omega}}_0 \underline{\theta}$$

and

$$\underline{\Omega}' = \underline{\dot{\theta}}$$

relation (11) leads to

$$\underline{\Omega} = {}^0\underline{\Omega}_0 + \delta \underline{\Omega} \quad (12.1)$$

with

$$\delta \underline{\Omega} = {}^0\underline{\tilde{\Omega}}_0 \underline{\theta} + \underline{\dot{\theta}} \quad (12.2)$$

3.3 Linearization of overall rotational equations

Substituting ${}^0\underline{\Omega}_0 + \delta \underline{\Omega}$ for $\underline{\Omega}$ in relation (9) yields, at first order :

$$(J_E + \sum_i I_n^i) \delta \underline{\dot{\Omega}} + {}^0\underline{\tilde{\Omega}}_0 (J_E + \sum_i I_n^i) - (J_E + \sum_i I_n^i) {}^0\underline{\Omega}_0 + \sum_i \dot{\theta}^i (\tilde{w}^i I_n^i - I_n^i \tilde{w}^i - I_n^i \tilde{w}^i) \delta \underline{\Omega} = \underline{\Gamma}_0 + \underline{\mathcal{H}}_{(ex)R}$$

where :

$$\begin{aligned} \underline{\Gamma}_0 = & -(J_E + \sum_i I_n^i) {}^0\underline{\dot{\Omega}}_0 \\ & - {}^0\underline{\tilde{\Omega}}_0 (J_E + \sum_i I_n^i) {}^0\underline{\Omega}_0 \\ & - \sum_i \dot{\theta}^i (\tilde{w}^i I_n^i - I_n^i \tilde{w}^i - I_n^i \tilde{w}^i) {}^0\underline{\Omega}_0 \\ & - \sum_i \dot{\theta}^i I_n^i \tilde{w}^i \\ & - \sum_i (\dot{\theta}^i)^2 \tilde{w}^i I_n^i \tilde{w}^i \end{aligned}$$

The torque $\underline{\mathcal{H}}_{(ex)R}$ is the summation of

- the outer-space disturbing torques, say T_R
- the control torques, say u_R

By splitting the control torques as follows :

$$u_R = u_{OR} + u_R$$

and selecting

$$u_{OR} = -\underline{\Gamma}_0$$

one may derive the linearized form of the overall rotational equations, say :

$$(J_E + \sum_i I_n^i) \delta \underline{\dot{\Omega}} + {}^0\underline{\tilde{\Omega}}_0 (J_E + \sum_i I_n^i) - (J_E + \sum_i I_n^i) {}^0\underline{\Omega}_0 + \sum_i \dot{\theta}^i (\tilde{w}^i I_n^i - I_n^i \tilde{w}^i - I_n^i \tilde{w}^i) \delta \underline{\Omega} = T_R + u_R \quad (13)$$

Now let the prescribed behavior be such that ${}^0\underline{\dot{\Omega}}_0 = 0$

and let \underline{X} be defined by the relation :

$$\underline{X} = (J_E + \sum_i I_n^i)^{-1} {}^0\underline{\tilde{\Omega}}_0 (J_E + \sum_i I_n^i) - (J_E + \sum_i I_n^i) {}^0\underline{\Omega}_0 + \sum_i \dot{\theta}^i (\tilde{w}^i I_n^i - I_n^i \tilde{w}^i - I_n^i \tilde{w}^i) \quad (14)$$

Substituting $({}^0\underline{\tilde{\Omega}}_0 \underline{\theta} + \underline{\dot{\theta}})$ for $\delta \underline{\Omega}$ in relation (13) finally yields the expression of the satellite's linearized equations of rotation :

$$\underline{\ddot{\theta}} = -[{}^0\underline{\tilde{\Omega}}_0 + \underline{X}] \underline{\dot{\theta}} - [\underline{X} {}^0\underline{\tilde{\Omega}}_0] \underline{\theta} + (J_E + \sum_i I_n^i)^{-1} T_R + (J_E + \sum_i I_n^i)^{-1} u_R \quad (15)$$

APPENDIX II.

LINEAR SYSTEMS AFFECTED BY DETERMINISTIC DISTURBANCES :
 •General results relative to observability properties•

This appendix aims at deriving an observability canonical form for time invariant linear systems the state-space representation of which may read as

$$(S_a) \quad \begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} & \begin{matrix} A_{11}(\ell, \ell) \quad , \quad A_{12}(\ell, m) \\ A_{22}(m, m) \end{matrix} \\ y = [C \quad 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{cases}$$

and which are endowed with the following properties

$$(H) \quad \begin{cases} a) (A_{11}, C) \text{ is completely observable} \\ b) \text{ matrices } A_{11} \text{ and } A_{22} \text{ have no common eigenvalue} \\ c) \text{ matrix } A_{22} \text{ is diagonalizable} \end{cases}$$

Such systems and properties will be respectively referred to as (S_a) and (H) hereinafter.

It is demonstrated that a geometric approach does away with computation of the observability matrix.

Furthermore the unobservability problems are split onto the different eigen-subspaces of matrix A_{22} .

1. Preliminary observability theorem

Theorem 1 : Let that time-invariant linear system (S') whose state-space representation reads as

$$(S') \quad \begin{cases} \dot{X} = A' X \\ Y = C' X \end{cases}$$

Now, let P'_m , defined as

$$P'_m(x) = \prod_{i=1}^{p'} (\lambda'_i - x)^{\beta'_i}$$

be the minimal polynomial of A' , and let N'_i be the characteristic subspace of A' relative to the eigenvalue λ'_i .

Then the unobservable subspace for the system (S') appears as

$$V' = \sum_{i=1}^{p'} V'_i$$

with :

$$V'_i = \{ N'_i \cap \text{Ker } C' \cap \text{Ker } C'A' \cap \dots \cap \text{Ker } C'A'^{(\beta'_i-1)} \}$$

Demonstration of that theorem may be found under reference.

2. Characterization of the unobservable subspace for system (S_a)

Such characterization is performed by setting

$$\begin{aligned} A' &= \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \\ C' &= [C \quad 0] \end{aligned}$$

and by straightforward application of the theorem here-above.

2.1 Notations

- $\sigma(A')$ is the spectrum of A'
- $\sigma(A_{11}) = \{\lambda'_i; 1 \leq i \leq p\}$ is the spectrum of A_{11}
- $\sigma(A_{22}) = \{\mu'_j; 1 \leq j \leq q\}$ is the spectrum of A_{22}
- $P'_m(x) = \prod_{i=1}^p (\lambda'_i - x)^{\beta'_i}$ is the minimal polynomial of A_{11}
- $N'_i = \text{Ker}(A_{11} - \lambda'_i I_p)^{\beta'_i}$ is the characteristic subspace of A_{11} relative to the eigenvalue λ'_i
- $E_j = \text{Ker}(A_{22} - \mu'_j I_m)$ is the eigensubspace of A_{22} relative to the eigenvalue μ'_j (cf. assumption (H.c)).

2.2 Characteristic subspaces of A'

Readily we have :

$$\sigma(A') = \sigma(A_{11}) \cup \sigma(A_{22})$$

Then let N'_i ($1 \leq i \leq p$) and N'_j ($1 \leq j \leq q$) be the characteristic subspaces of A' relative to the eigenvalues λ'_i and μ'_j respectively.

These subspaces can easily be demonstrated, since A_{11} and A_{22} have no common eigenvalue (cf. assumption (H.b)), to appear as :

$$N'_i = \{(x_1, x_2) : x_1 \in N_i \text{ and } x_2 = 0\} \quad 1 \leq i \leq p$$

$$N'_j = \{(x_1, x_2) : x_1 = -F_j x_2 \text{ and } x_2 \in E_j\} \quad 1 \leq j \leq q$$

with :

$$F_j = (A_{11} - \mu'_j I_p)^{-1} A_{12}$$

Correlatively, the minimal polynomial P'_m of A' is expressed as :

$$P'_m(x) = \prod_{i=1}^p (\lambda'_i - x)^{\beta'_i} \prod_{j=1}^q (\mu'_j - x)$$

which implies that the coefficient β_i' :

- equals β_i if λ_i' belongs to $\sigma(A_{11})$
- equals 1 if λ_i' belongs to $\sigma(A_{22})$

2.3 Unobservable subspace for system (S_a)

Let :

$$\begin{aligned} V_1^1 &= \{ N_1^1 \cap \text{Ker } C' \cap \text{Ker } C' A' \cap \dots \cap \text{Ker } C' A'^{(\beta_1-1)} \} \quad 1 \leq i \leq p \\ V_j^2 &= \{ N_j^2 \cap \text{Ker } C' \} \quad 1 \leq j \leq q \end{aligned}$$

Then obviously :

$$(x_1, x_2) \in V_1^1 \Leftrightarrow x_1 \in \{ N_1 \cap \text{Ker } C \cap \dots \cap \text{Ker } C A_{11}^{(\beta_1-1)} \} \quad \text{and} \quad x_2 = 0$$

Since (A_{11}, C) is, by assumption (H.a), completely observable, the theorem in §1 provides the demonstration that

$$N_1 \cap \text{Ker } C \cap \dots \cap \text{Ker } C A_{11}^{(\beta_1-1)} = 0$$

which results in the affirmation that :

$$V_1^1 = 0$$

Similarly :

$$\begin{aligned} (x_1, x_2) \in V_j^2 &\Leftrightarrow \begin{cases} x_1 = -F_j x_2 \\ C F_j x_2 = 0 \\ x_2 \in E_j \end{cases} \\ &\Leftrightarrow \{ x_1 = -F_j x_2 \quad \text{and} \quad x_2 \in [E_j \cap \text{Ker } C F_j] \} \end{aligned}$$

The following theorem can now be spelt out :

Theorem 2 : Let matrix F_j and subspace V_j be respectively defined as follows :

$$\begin{aligned} F_j &= (A_{11} - \mu_j I_q)^{-1} A_{12} \quad 1 \leq j \leq q \\ V_j &= \{ (x_1, x_2) : x_1 = -F_j x_2 ; x_2 \in E_j \cap \text{Ker } C F_j \} \quad 1 \leq j \leq q \end{aligned}$$

Then the unobservable subspace for the system (S_a) appears as :

$$\mathcal{V} = \bigoplus_{j=1}^q V_j$$

3. Derivation of an observability canonical form for system (S_a)

3.1 Reminder

Let E^* be the dual of E . Let \mathcal{V}^1 be the annihilator of the unobservable subspace \mathcal{V} :

$$\mathcal{V}^1 = \{ \varphi \in E^* ; \varphi(x) = 0 \quad \forall x \in \mathcal{V} \}$$

Let any element of the dual space be expressed by a row vector such that \mathcal{V}^1 may read as any matrix N^T with independent rows and satisfying the equality :

$$N^T x = 0 \quad \forall x \in \mathcal{V}$$

Then an observability canonical form is obtained by setting

$$X = N^T x.$$

3.2 Characterization of the annihilator \mathcal{V}^1 of \mathcal{V}

Let the eigenvalues μ_j be ordered in such a way that :

$$\begin{aligned} V_j &\neq 0 \quad \text{if} \quad j \leq K \\ V_j &= 0 \quad \text{if} \quad K < j \leq q. \end{aligned}$$

Then theorem 2 gives us :

$$\mathcal{V} = \bigoplus_{j=1}^K V_j$$

Now, let Q_j be the projection operator on E_j along the direct sum $\bigoplus_{i \neq j} E_i$.

The following is then readily demonstrated :

$$(x_1, x_2) \in \mathcal{V} \Leftrightarrow \begin{cases} x_1 = - \sum_{j=1}^K F_j Q_j x_2 \quad \text{and} \quad x_2 \in \sum_{j=1}^K (E_j \cap \text{Ker } C F_j) \end{cases}$$

since :

$$x_2 \in \sum_{j=1}^K (E_j \cap \text{Ker } C F_j) \Leftrightarrow \begin{cases} x_2 \in \sum_{j=1}^K E_j \quad \text{and} \quad (Q_j x_2 \in \text{Ker } C F_j \quad \forall j \quad 1 \leq j \leq K) \end{cases}$$

$$\Leftrightarrow \begin{cases} (I_m - \sum_{j=1}^K Q_j) x_2 = 0 \\ C F_j Q_j x_2 = 0 \quad \forall j \quad 1 \leq j \leq K \end{cases}$$

we finally obtain

$$(x_1, x_2) \in \mathcal{V} \Leftrightarrow \begin{bmatrix} I_q & \sum_{j=1}^K F_j Q_j \\ 0 & I_m - \sum_{j=1}^K Q_j \\ 0 & C F_1 Q_1 \\ \vdots & \vdots \\ 0 & C F_K Q_K \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

Then, the annihilator \mathcal{V}^\perp of \mathcal{V} is characterized by the independent rows of the matrix hereabove.

3.3 An observability canonical form for system (S_a)

Let S_0^T and S_j^T ($1 \leq j \leq K$) be respectively the matrices built with the independent rows of the matrices $(I_m - \sum_{j=1}^K Q_j)$ and $C F_j Q_j$ ($1 \leq j \leq K$).

Now we shall demonstrate the linear independence of those subspaces of E^* which are respectively generated by the S_j^T ($0 \leq j \leq K$) rows.

Let s_j^* be any element of S_j^T ($0 \leq j \leq K$).

Then, the relation :

$$\sum_{j=0}^K a_j s_j^* = 0$$

reads :

$$\sum_{j=0}^K a_j s_j^*(x) = 0 \quad \forall x \in E$$

Now, since for $x \in E_i$, we have :

$$s_j(x) = 0 \quad \forall j \neq i,$$

we derive from the preceding relation :

$$a_i s_i^*(x) = 0 \quad \forall x \in E_i \quad 1 \leq i \leq K$$

i.e. :

$$a_i = 0 \quad 1 \leq i \leq K$$

Finally, we obtain :

$$a_j = 0 \quad 0 \leq j \leq K$$

We can now set the following theorem :

Theorem 3 : With the above notations, an observability canonical form for system (S_a) may read as :

$$\begin{cases} x_1 = x_1 + \left\{ \sum_{j=1}^K F_j Q_j \right\} x_2 \\ x_2 = \begin{bmatrix} S_0^T \\ S_1^T \\ \vdots \\ S_K^T \end{bmatrix} x_2 \end{cases}$$

4. Derivation of an optimal observability canonical form for system (S_a)

We shall confine ourselves to the case of periodic disturbances.

The eigenvalue of matrix A_{22} can therefore be written :

$$\mu_j = j\omega \quad \bar{\mu}_j = -j\omega$$

We set

$$G_j^T = F_j Q_j \quad 1 \leq j \leq K$$

$$G^T = \sum_{j=1}^K G_j^T$$

and we denote by 1A and 2A respectively the real and imaginary parts of the complex matrix A .

With the above notations, the observability canonical form derived in theorem 3 (see section 3) may read as:

$$\begin{cases} x_1 = x_1 + (G^T + \bar{G}^T) x_2 \\ x_{2j} = S_j^T x_2 \quad 0 \leq j \leq K \end{cases}$$

Now, let

$$\sum_{j=1}^K \Lambda_j^T S_j^T = G^T - \epsilon^T$$

be any projection in E^* of G^T onto the subspaces S_j^T .

Then, any form of the type

$$\begin{cases} x_1 = x_1 + (\epsilon^T + \bar{\epsilon}^T) x_2 \\ x_{2j} = S_j^T x_2 \quad 0 \leq j \leq K \end{cases}$$

will be canonical for observability.

Our purpose is to utilize a priori knowledge of the disturbances to calculate the Λ_j^T matrices in such a way as to minimize the difference $X_{1i} - \hat{X}_{1i}$ ($1 \leq i \leq \ell$), thus leading to the criterion:

$$J_i = E \left\{ \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \|X_{1i} - \hat{X}_{1i}\| dt \right\}$$

where $E(\cdot)$ denotes the mathematical expectation.

After a few manipulations, matrix Λ_j^T may be shown as :

$$\Lambda_j^T = (S_j)_{M_j}^* G_j$$

where semi-norm M_j is expressed as :

$$M_j = E \left\{ \frac{\omega_j}{2\pi} \int_0^{2\pi/\omega_j} |Q_j x_2| |Q_j x_2|^T dt \right\}$$

and where

$$(S_j)_{M_j}^* = (S_j^* M_j S_j)^* S_j^* M_j$$

denotes the pseudo inverse of the matrix S_j for the semi-norm M_j .

Now, we proceed with the computation of $\dot{\hat{X}}_1$.

Keeping in mind that :

$$\frac{d}{dt}(G_j x_2) = \mu_j G_j x_2$$

leads to :

$$\frac{d}{dt}(e^{\mu_j t} x_2) = \sum_j \mu_j (G_j^T - \Lambda_j^T S_j^T) x_2$$

Hence

$$\begin{aligned} \dot{\hat{X}}_1 = A_{11} X_1 + A_{12} x_2 - \sum_j (A_{11} - \mu_j I_{\ell}) (G_j^T - \Lambda_j^T S_j^T) x_2 \\ - \sum_j (A_{11} + \mu_j I_{\ell}) (\bar{G}_j^T - \bar{\Lambda}_j^T \bar{S}_j^T) x_2 \end{aligned}$$

Since

$$\begin{aligned} (A_{11} - \mu_j I_{\ell}) G_j^T &= (A_{11} - \mu_j I_{\ell}) F_j Q_j \\ &= A_{12} Q_j \end{aligned}$$

the following identity holds :

$$(A_{11} - \mu_j I_{\ell}) (G_j^T - \Lambda_j^T S_j^T) = A_{12} Q_j - A_{12} Q_j [(S_j)_{M_j}^*]^T S_j^T$$

Finally, we have :

$$\begin{aligned} \dot{\hat{X}}_1 = A_{11} X_1 + A_{12} \left\{ I_m - \sum_j (Q_j + \bar{Q}_j) \right\} x_2 + A_{12} \left\{ \sum_j Q_j [(S_j)_{M_j}^*]^T S_j^T \right. \\ \left. + A_{12} \left\{ \sum_j \bar{Q}_j (\bar{S}_j)_{\bar{M}_j}^* \right\}^T \bar{S}_j^T \right\} x_2 \end{aligned}$$

Considering the isomorphism of algebra

$$\mathcal{M}_n(\mathbb{C}) \rightarrow \mathcal{M}_{2n}(\mathbb{R})$$

$$A = {}^1A + j {}^2A \mapsto f(A) = \begin{bmatrix} {}^1A & -{}^2A \\ {}^2A & {}^1A \end{bmatrix}$$

where $\mathcal{M}_n(K)$ denotes the algebra of square matrices $n \times n$ on the field K , gives us :

Theorem 4 : The optimal observability canonical form for system (S_j) appears as :

$$\begin{cases} \dot{X}_1 = x_1 + 2 \sum_j N_j x_2 \\ \dot{X}_{20} = S_0^T x_2 \\ {}^1\dot{X}_{2j} = {}^1S_j^T x_2 \\ {}^2\dot{X}_{2j} = {}^2S_j^T x_2 \end{cases}$$

with

$$N_j = {}^1G_j^T - [{}^1G_j^T \mid {}^2G_j^T] [(S_j)_{M_j}^*]^T \begin{bmatrix} {}^1S_j^T \\ {}^2S_j^T \end{bmatrix}$$

$$M_j^* = \begin{bmatrix} M_j & 0 \\ 0 & M_j \end{bmatrix}$$

$$S_j^* = \begin{bmatrix} {}^1S_j^T & -{}^2S_j^T \\ {}^2S_j^T & {}^1S_j^T \end{bmatrix}$$

Theorem 5 : The state equations of the optimal canonical form may read as :

$$\begin{cases} \dot{\hat{X}}_1 = A_{11} X_1 + A_{12} B X_{20} + 2A_{12} \left\{ \sum_j [{}^1Q_j \mid {}^2Q_j] [(S_j)_{M_j}^*]^T \right\} \begin{bmatrix} {}^1x_{2j} \\ -{}^2x_{2j} \end{bmatrix} \\ \dot{\hat{X}}_{20} = A_{22}^* X_{20} \\ {}^1\dot{\hat{X}}_{2j} = -\omega_j {}^2x_{2j} \\ {}^2\dot{\hat{X}}_{2j} = \omega_j {}^1x_{2j} \end{cases}$$

where B is given by

$$[I_m - 2 \sum_j {}^1Q_j] = B S_0^T$$

and A_{22}^* is the matrix of the restriction of A_{22} to the subspace $\sum_{j=1}^{\ell} E_j$.

APPENDIX III.

LITERAL CALCULATIONS

This appendix develops all literal calculations required to obtain the linearized dynamic equations of the spacecraft as well as the system's optimal canonical form for observability. It directly applies the results of the foregoing two appendices to the specific satellite case considered in the paper.

1. Spacecraft linearized dynamic equations

1.1 Spacecraft configuration

The satellite consists of

- one assumedly rigid, solid body S
- one momentum wheel S¹
- assumedly rigid solar array wings S²

The wheel S¹ and the panels S² rotate about axis \vec{V} of frame \mathcal{A} .
Transition matrix P² from frame \mathcal{A} to frame \mathcal{A}^2 reads:

$$P^2 = \begin{bmatrix} \cos \theta_2 & 0 & \sin \theta_2 \\ 0 & 0 & 0 \\ -\sin \theta_2 & 0 & \cos \theta_2 \end{bmatrix}$$

As a simplification, the centers of gravity G, G¹ and G² are assumed to be aligned.

Let G_e be the center of gravity of the whole spacecraft.

The vectors

$$\begin{aligned} \vec{G_E G} &= \vec{d} - \vec{d_e} \\ \vec{G_E G^1} &= \vec{d^1} - \vec{d_e} \\ \vec{G_E G^2} &= \vec{d^2} - \vec{d_e} \end{aligned}$$

and the inertia tensors

$$J, J^1, J^2$$

are selected as equal, respectively, to :

$$\begin{aligned} G_E G & \begin{vmatrix} 0 \\ d - d_e \\ 0 \end{vmatrix} & G_E G^1 & \begin{vmatrix} 0 \\ d^1 - d_e \\ 0 \end{vmatrix} & G_E G^2 & \begin{vmatrix} 0 \\ 0 \\ 0 \end{vmatrix} \\ \text{for the vectors,} & & & & & \\ \text{and} & J = \begin{bmatrix} I_1 & & \\ & I^2 & \\ & & I_3 \end{bmatrix} & J^1 = \begin{bmatrix} h_x & & \\ & h_y & \\ & & h_x \end{bmatrix} & J^2 = \begin{bmatrix} I_{p_x} & & \\ & I_{p_y} & \\ & & I_{p_z} \end{bmatrix} \\ \text{for the inertia tensors.} & & & & & \end{aligned}$$

1.2 Prescribed behavior

First, with the classical conventions, the instant local orbital frame

$\mathcal{A}_L = (G_E, \vec{U}_L, \vec{V}_L, \vec{W}_L)$ is defined :

- \vec{W}_L : yaw, along the downward geocentric
- \vec{U}_L : roll, within orbit plane, such that $\vec{U}_L \perp \vec{W}_L$, with \vec{U}_L oriented in the same sense as velocity of G_E
- \vec{V}_L : pitch such that \mathcal{A}_L is a direct orthonormed frame.

Relative to inertial frame \mathcal{A} , frame \mathcal{A}_L rotates about \vec{V}_L with angular velocity $-\omega$ (ω modulus of earth rotation velocity).

Spacecraft frame \mathcal{A} is the direct orthonormed frame formed by the principal inertia axes of main body S. If the prescribed behavior is fulfilled, spacecraft frame \mathcal{A} and local orbital frame \mathcal{A}_L are coincident at all time, i.e.

$$\mathcal{A}_0 = \mathcal{A}_L$$

The three pseudo Euler angles that enable transition from frame \mathcal{A}_L to frame \mathcal{A} appear respectively as :

$$\underline{\theta} = [\varphi, \theta, \psi]^T$$

with

$$\begin{aligned} \varphi &= \text{roll} \\ \theta &= \text{pitch} \\ \psi &= \text{yaw} \end{aligned}$$

Then, for the prescribed behavior :

$$\begin{aligned} {}^0\Omega_0 &= [0 \quad -\omega \quad 0]^T \\ \dot{\theta}_1 &= \omega_1 \\ \dot{\theta}_2 &= \omega \end{aligned}$$

1.3 Computation of $(J_E + \Sigma \text{In}^i)$ and X (see appendix 1, section 3.3)

Let

$$\begin{aligned} a(t) &= (I_1 + md^2) + (h_x + m_1(d_1^2) + \frac{1}{2}[(I_{p_x} + I_{p_z}) - (I_{p_z} - I_{p_x}) \cos(2\theta_2)]) \\ b(t) &= \frac{I_{p_z} - I_{p_x}}{2} \sin(2\theta_p) \\ c(t) &= I_2 - (I_3 + md^2) + (h_y - h_x - m_1(d^1)^2) - \frac{1}{2}[(I_{p_x} + I_{p_z}) - (I_{p_z} - I_{p_x}) \cos(2\theta_2)] \\ e(t) &= (I_3 + md^2) + (h_x + m_1(d^1)^2) + \frac{1}{2}[(I_{p_x} + I_{p_z}) + (I_{p_z} - I_{p_x}) \cos(2\theta_2)] \\ f(t) &= (I_1 + md^2 - I_2) + (h_x + m_1(d^1)^2 - h_y) + \frac{1}{2}[(I_{p_x} + I_{p_z}) + (I_{p_z} - I_{p_x}) \cos(2\theta_2)] \\ g &= I_2 + h_y + I_{p_y} \\ H &= h_y \omega_1 \end{aligned}$$

Then :

$$J_E = \begin{bmatrix} a & 0 & b \\ 0 & g & 0 \\ b & 0 & e \end{bmatrix}$$

$$J_E^{-1} = \begin{bmatrix} \frac{e}{ae-b^2} & 0 & \frac{-b}{ae-b^2} \\ 0 & \frac{1}{g} & 0 \\ \frac{-b}{ae-b^2} & 0 & \frac{a}{ae-b^2} \end{bmatrix}$$

and :

$$X = \frac{1}{ae-b^2} \begin{bmatrix} -b[\omega(f-e) + H] & 0 & -[-\omega(b^2+ce) + He] \\ 0 & 0 & 0 \\ -[\omega(b^2-af) - HA] & 0 & -b[\omega(a+c) - H] \end{bmatrix}$$

1.4 Linearized roll-yaw equations

Let T_Z and T_X be respectively the components on the yaw and roll axis of the disturbing torques which affect the satellite.

Setting

$$x = \frac{He}{ae-b^2}$$

$$y = \frac{Hb}{ae-b^2}$$

$$z = \frac{Ha}{ae-b^2}$$

provides the linearized roll-yaw equations with the form :

$$\begin{cases} \ddot{\phi} = (\omega_x)\dot{\phi} + y\ddot{\psi} - (\omega_y)\dot{\psi} + x\ddot{\psi} - \frac{y}{H}T_Z + \frac{x}{H}T_X \\ \ddot{\psi} = -(\omega_y)\dot{\phi} - z\ddot{\phi} + (\omega_z)\dot{\psi} - y\ddot{\phi} + \frac{z}{H}T_Z - \frac{y}{H}T_X \end{cases}$$

1.5 Numerical values

The following values were adopted in all numerical applications :

$$I_1 = I_2 = I_3 = 1,000 \text{ Kg.m}^2$$

$$h_x \cong 0$$

$$h_y = 0.9625 \text{ Kg.m}^2$$

$$I_{PX} = 300 \text{ Kg.m}^2 \quad I_{PY} = 4,300 \text{ Kg.m}^2 \quad I_{PZ} = 4,000 \text{ Kg.m}^2$$

$$H = h_y \omega_1 = 50 \text{ Nm.s}$$

$$\omega = 7.27 \times 10^{-5} \text{ rad/s.}$$

2. Optimal canonical form for observability

Matrix A_{22} has three double eigenvalues, which are

$$\mu_1 = 0 \quad \mu_2 = j\omega \quad \mu_3 = -j\omega$$

We shall denote the three relative eigensubspaces by E_1, E_2 and \bar{E}_2 respectively.

Matrix Q_1 of projection on eigensubspace E_1 reads :

$$Q_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Matrix Q_2 of projection in eigensubspace E_2 reads :

$$Q_2 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & j\omega & 0 & 0 \\ 0 & 0 & j\omega & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & j\omega \\ 0 & 0 & 0 & 0 & j\omega & 1 \end{bmatrix}$$

Straightforwardly, the projection matrix on eigensubspace \bar{E}_2 reads as \bar{Q}_2 .

Computation of matrix $(A_{11} - pI_4)^{-1}$ by Leverrier - Souriau's algorithm yields :

$$(A_{11} - pI_4)^{-1} = -\frac{B_0 p^3 + B_1 p^2 + B_2 p + B_3}{d_0 p^4 + d_1 p^3 + d_2 p^2 + d_3 p + d_4}$$

with :

$$B_0 = I_4 \quad B_1 = A_{11}$$

$$B_2 = \begin{bmatrix} xz - y^2 - \omega z & y & -\omega y & x \\ 0 & -\omega z & -\omega(y^2 - xz) & -\omega y \\ -\omega y & -z & xz - y^2 - \omega x & -y \\ \omega(y^2 - xz) & -\omega y & 0 & -\omega x \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 0 & -\omega z & -\omega(y^2 - xz) & -\omega y \\ \omega^2(y^2 - xz) & 0 & 0 & 0 \\ \omega(y^2 - xz) & -\omega y & 0 & -\omega x \\ 0 & 0 & \omega^2(y^2 - xz) & 0 \end{bmatrix}$$

2.1 Computation of G_1^T and S_1^T

Knowing that

$$G_1^T = F_1 Q_1$$

$$S_1^T = C F_1 Q_1$$

with

$$F_1 = A_{11}^{-1} A_{12}$$

we readily obtain

$$G_1^T = \begin{bmatrix} 0 & \frac{1}{H\omega} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{H\omega} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and :

$$S_1^T = C F_1 Q_1 = \begin{bmatrix} 0 & \frac{1}{H\omega} & 0 & 0 & 0 & 0 \end{bmatrix}$$

2.2 Computation of ${}^1G_2^T$, ${}^2G_2^T$, ${}^1S_2^T$, ${}^2S_2^T$

Setting :

$$A = y \omega$$

$$B = y^2 - x z$$

$$C = x \omega + (x z - y^2)$$

$$D = z \omega + (x z - y^2)$$

$$E = \omega^2 (\omega + x + z)$$

and knowing that :

$${}^1G_2^T = \mathcal{R}e(F_2 Q_2)$$

$${}^2G_2^T = \mathcal{I}m(F_2 Q_2)$$

$${}^1S_2^T = \mathcal{R}e(C F_2 Q_2)$$

$${}^2S_2^T = \mathcal{I}m(C F_2 Q_2)$$

with :

$$F_2 = (A_{11} - j\omega L_4)^{-1} A_{12}$$

we readily obtain :

$${}^1G_2^T = \frac{1}{2HE} \begin{bmatrix} 0 & 0 & -A\omega & B & C\omega & 0 \\ 0 & 0 & -B\omega^2 & -A\omega & 0 & C\omega \\ 0 & 0 & D\omega & 0 & -A\omega & -B \\ 0 & 0 & 0 & D\omega & B\omega^2 & -A\omega \end{bmatrix}$$

$${}^2G_2^T = \frac{1}{2HE} \begin{bmatrix} 0 & 0 & B\omega & A & 0 & -C \\ 0 & 0 & -A\omega^2 & B\omega & C\omega^2 & 0 \\ 0 & 0 & 0 & -D & -B\omega & A \\ 0 & 0 & D\omega^2 & 0 & -A\omega^2 & -B\omega \end{bmatrix}$$

$${}^1S_2^T = \frac{1}{2HE} \begin{bmatrix} 0 & 0 & -A\omega & B & C\omega & 0 \end{bmatrix}$$

$${}^2S_2^T = \frac{1}{2HE} \begin{bmatrix} 0 & 0 & B\omega & A & 0 & -C \end{bmatrix}$$

2.3 Computation of $\{(S_i^T)_{M_i}^*\}^T$

Assuming the expectations of the amplitudes of sine disturbing torques to be known, i.e.

$$E\{|T_{Z_1}|^2\} = \gamma^2$$

$$E\{|T_{X_1}|^2\} = \delta^2$$

yet with no assumptions as to their phases (probability uniform on $[0, 2\pi]$), matrix M_2 :

$$M_2 = \frac{\omega}{2\pi} E \left\{ \int_0^{2\pi} [{}^1Q_2 x_2] [{}^1Q_2 x_2]^T dt \right\}$$

may read as :

$$M_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \gamma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega^2 \gamma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \omega^2 \delta^2 \end{bmatrix}$$

Now :

$$(S_i^T)_{M_i}^* = (S_i^T M_i^T S_i^T)^* S_i^T M_i^T$$

We shall first compute the term $(S_i^T M_i^T S_i^T)^*$.

We have :

$$S_i^T M_i^T S_i^T = \begin{bmatrix} {}^1S_2^T & {}^2S_2^T \\ -{}^2S_2^T & {}^1S_2^T \end{bmatrix} \begin{bmatrix} M_2 & 0 \\ 0 & M_2 \end{bmatrix} \begin{bmatrix} {}^1S_2 & -{}^2S_2 \\ {}^2S_2 & {}^1S_2 \end{bmatrix}$$

Hence :

$$S_i^T M_i^T S_i^T = \frac{2\omega^2 \{ (A^2 + B^2) \gamma^2 + C^2 \delta^2 \}}{(2HE)^2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$(S_2^T \ M_2^T \ S_2^T)^* = \frac{(2HE)^2}{2\omega^2 \{(A^2+B^2)\gamma^2 + C^2 \delta^2\}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thereby :

$$\{(S_2^T \ M_2^T \ S_2^T)^*\}^T = \frac{(2HE)^2}{2\omega^2 \{(A^2+B^2)\gamma^2 + C^2 \delta^2\}} \begin{bmatrix} M_2 & 0 \\ 0 & M_2 \end{bmatrix} \begin{bmatrix} {}^1S_2 & -{}^2S_2 \\ {}^2S_2 & {}^1S_2 \end{bmatrix}$$

2.4 Computation of $[{}^1\Lambda_2^T \mid {}^2\Lambda_2^T]$

Remembering that

$$[{}^1\Lambda_2^T \mid {}^2\Lambda_2^T] = [{}^1G_2^T \mid {}^2G_2^T] \{(S_2^T \ M_2^T \ S_2^T)^*\}^T$$

gives us

$${}^1\Lambda_2^T = {}^1G_2^T M_2 {}^1S_2 + {}^2G_2^T M_2 {}^2S_2$$

$${}^2\Lambda_2^T = -{}^1G_2^T M_2 {}^2S_2 + {}^2G_2^T M_2 {}^1S_2$$

whence :

$$[{}^1\Lambda_2^T \mid {}^2\Lambda_2^T] = \begin{bmatrix} 1 & 0 \\ 0 & \omega \\ \frac{-A(D\gamma^2 + C\delta^2)}{(A^2+B^2)\gamma^2 + C^2\delta^2} & \frac{-B(D\gamma^2 + C\delta^2)}{(A^2+B^2)\gamma^2 + C^2\delta^2} \\ \frac{B\omega(D\gamma^2 + C\delta^2)}{(A^2+B^2)\gamma^2 + C^2\delta^2} & \frac{-A\omega(D\gamma^2 + C\delta^2)}{(A^2+B^2)\gamma^2 + C^2\delta^2} \end{bmatrix}$$

2.5 Computation of $(2N_2)$

Having in mind that

$$N_2 = {}^1G_2^T - [{}^1\Lambda_2^T \mid {}^2\Lambda_2^T] \begin{bmatrix} {}^1S_2^T \\ -{}^2S_2^T \end{bmatrix}$$

and setting :

$$X = D\gamma^2 + C\delta^2 \\ Z = (A^2+B^2)\gamma^2 + C^2\delta^2$$

we obtain :

$$2N_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \{D - \frac{(A^2+B^2)X}{Z}\}\omega & 0 & \{\frac{CX}{Z} - 1\}A\omega & \{\frac{CX}{Z} - 1\}B \\ 0 & 0 & 0 & \{D - \frac{(A^2+B^2)X}{Z}\}\omega & \{1 - \frac{CX}{Z}\}B\omega^2 & \{\frac{CX}{Z} - 1\}A\omega \end{bmatrix}$$

2.6 Computation of $2A_{12} [{}^1Q_2 \mid {}^2Q_2] \{(S_2^T \ M_2^T \ S_2^T)^*\}^T$

The equality

$$[{}^1Q_2 \mid {}^2Q_2] \begin{bmatrix} M_2 & 0 \\ 0 & M_2 \end{bmatrix} \begin{bmatrix} {}^1S_2 & -{}^2S_2 \\ {}^2S_2 & {}^1S_2 \end{bmatrix} = \frac{1}{2HE} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -A\gamma^2\omega & B\gamma^2\omega^2 \\ B\gamma^2\omega^2 & C\delta^2\omega \\ 0 & 0 \\ 0 & C\delta^2\omega^2 \end{bmatrix}$$

leads to :

$$2A_{12} [{}^1Q_2 \mid {}^2Q_2] \{(S_2^T \ M_2^T \ S_2^T)^*\}^T = \frac{2E}{\omega^2 \{(A^2+B^2)\gamma^2 + C^2 \delta^2\}} \begin{bmatrix} 0 & 0 \\ (yA\gamma^2 + xC\delta^2)\omega & yB\gamma^2\omega \\ 0 & 0 \\ -(zA\gamma^2 + yC\delta^2)\omega & -zB\gamma^2\omega \end{bmatrix}$$

2.7 Computation of Λ_1^T and N_1

Straightforwardly :

$$\Lambda_1^T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

and

$$N_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{H\omega} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

2.8 Optimal canonical form for observability

This may read as

$$\begin{cases} X_1 = x_1 + (N_1 + 2N_2)x_2 \\ X_{21} = S_1^T x_2 \\ {}^1X_{22} = {}^1S_2^T x_2 \\ {}^2X_{22} = {}^2S_2^T x_2 \end{cases}$$

with (see sections 2.5 and 2.7)

$$N_1 + 2N_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{H\omega} & 0 & \{D - \frac{A^2+B^2}{Z}\}X & 0 & \{\frac{CX}{Z} - 1\}A\omega & \{\frac{CX}{Z} - 1\}B \\ 0 & 0 & 0 & \{D - \frac{(A^2+B^2)X}{Z}\}\omega & \{1 - \frac{CX}{Z}\}B\omega^2 & \{\frac{CX}{Z} - 1\}A\omega \end{bmatrix}$$

and with (see sections 2.1 and 2.2) :

$$\begin{aligned} S_1^T &= \frac{1}{H\omega} [0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0] \\ {}^1S_2^T &= \frac{1}{2HE} [0 \quad 0 \quad -A\omega \quad B \quad C\omega \quad 0] \\ {}^2S_2^T &= \frac{1}{2HE} [0 \quad 0 \quad B\omega \quad A \quad 0 \quad -C] \end{aligned}$$

2.9 State equations for the optimal observability canonical form

For their deterministic part, and using the results of section 2.6, these equations appear as :

$$\begin{cases} \dot{X}_1 = A_{11} X_1 + \begin{bmatrix} 0 \\ \frac{x}{H} \\ 0 \\ -\frac{y}{H} \end{bmatrix} X_{21} + \frac{2E}{\omega^2 \{(A^2+B^2)\gamma^2 + C^2 \delta^2\}} \begin{bmatrix} 0 & 0 \\ (yA\gamma^2 + xC\delta^2)\omega & -yB\gamma^2\omega \\ 0 & 0 \\ -(zA\gamma^2 + yC\delta^2)\omega & zB\gamma^2\omega \end{bmatrix} \begin{bmatrix} {}^1X_{22} \\ {}^2X_{22} \end{bmatrix} \\ \dot{X}_{21} = 0 \\ {}^1\dot{X}_{22} = -\omega {}^2X_{22} \\ {}^2\dot{X}_{22} = \omega {}^1X_{22} \end{cases}$$

MULTIFUNCTION SPACECRAFT ATTITUDE ESTIMATION AND NAVIGATION SYSTEM

by

Dr. J. C. AMIEUX, B. CLAUDINON

MATRA Space Branch, VELIZY

ABSTRACT

The primary function of a spacecraft attitude control subsystem is the attitude determination and, more generally, the state estimation (attitude of the main body, appendages and flexible modes). The so-called "optico-inertial concept" is first described with application to a number of modern spacecraft; an example of implementation using space-qualified microprocessors is given in detail; the state estimation of a flexible spacecraft is then considered, a technique which can be readily implemented on existing hardware.

The extension of this concept to autonomous orbit control of an orbiting spacecraft is then considered for future development.

INTRODUCTION

The attitude control requirements of new coming spacecraft on the one hand, the qualification for space of inertial class gyroscopes on the other hand, have led the designer to implement the so-called optico-inertial concept. It is based on the optimal mixing of sensors having different spectral characteristics: gyroscopes with a high bandwidth low noise output, corrugated by long-term drift, and optical sensors with comparatively higher noise, lower bandwidth but no drift.

The technique implements basically the Kalman filter theory, which has been made possible by the qualification of spaceborn computers and even, as will be shown, by microprocessors.

The applicability of the concept seems only limited by its relatively high cost (development cost but also weight, reliability and power requirements on the spacecraft). As a result, this approach is used whenever required by the mission. As examples, and historically, the first missions to implement the technique were the astronomical sources observation (OAO, IUE, EXOSAT, and now HIPPARCOS). The second missions requiring high angular rate stability are the Earth observation missions (LANDSAT, SPOT). The third class of missions is the telecommunication mission from the geostationary orbit, especially when large solar arrays are implemented to provide the high power required by the payload, inducing large perturbing torques, incompatible with the more commonly used techniques based on angular momentum conservation.

Now, in order to increase the ratio of the payload mass to the service platform mass, light structural elements are used resulting in very flexible spacecraft, especially for the solar arrays but also large antennae and booms. A spacecraft state estimation is then required; it is defined as the main body angular position and rates, steerable appendages positions and rates, and flexible modes. The estimation of these variables is obtained by digital processing of high-bandwidth detectors outputs, (gyroscopes, relative motion detectors, infrared and radio frequency sensors).

The MATRA Space Branch has been involved in the development of this technique since 1977, in various research and development programs with C.N.E.S., E.S.A. and INTELSAT (SMAOII, SIOS, IASS and IADES contracts), and in spacecraft attitude control development for C.N.E.S. and E.S.A. (SPOT and HIPPARCOS).

The research and development programs now being conducted are oriented towards the rendez-vous and docking techniques between a chaser satellite and a target satellite, both in low, transfer and geostationary orbits. As a candidate for these applications, we are now thinking of increasing the autonomy of the orbiting spacecraft for navigation and orbit control function. This leads to the "autonomous station keeping" of a telecommunication satellite and, more generally, to autonomous spacecraft navigation as an extension of the basic concept.

OPTICO-INERTIAL CONCEPT AND ITS GENERALIZATION TO SPACECRAFT STATE ESTIMATION

As shortly exposed above, the so-called "optico-inertial concept" is based on optimal mixing of sensors outputs with different spectral characteristics.

The information gained on a number of flight-qualified sensors (infrared Earth sensors, Sun sensors, star scanners, star trackers, radiofrequency sensors, gyroscopes, accelerometers and resolvers) allows to define two broad classes of output spectra.

Figure 1 shows the output spectrum of most optical sensors, modelled as a colored white noise (i.e. white noise filtered at some cut-off frequency and sampled). Figure 2 shows the output spectrum of most inertial components (gyroscopes and accelerometers), modelled as a random walk process (low frequency) and a white noise filtered at some frequency (usually 5 to 10 Hz).

The mixing of these pieces of information is realized by an on-board Kalman filter, implemented in the software of the on-board computer. The optimal filtering approach is efficient in this application owing to the fact that both the dynamical model of the spacecraft and the sensors noise spectra are accurate, a rarely encountered situation.

AD-P003 387

1. GENERAL KINEMATIC MODELIZATION- Gyroscope model

The S/C rate on gyro sensitive axis is :

$$\underline{\omega} = \underline{\omega}_m + \underline{D} + \underline{b}_1$$

where \underline{b}_1 is a white centered noise

$$E(\underline{b}_1 \cdot \underline{b}_1^T) = Q_1 \delta(t-\tau)$$

\underline{D} is the gyro drift modelled by

$$\dot{\underline{D}} = \underline{b}_2 \text{ (random walk process)}$$

$$\underline{b}_2 = \text{white centered noise : } E(\underline{b}_2 \cdot \underline{b}_2^T) = Q_2 \delta(t-\tau)$$

$$\underline{\omega}_m = \text{gyro rate output.}$$

- Spacecraft angular rates model

When a redundant set of gyros is used $\underline{\omega}_m = M \cdot \underline{\Omega}$

where $\underline{\Omega}$ is the absolute rotation vector in body axes.

Therefore we obtain :

$$\underline{\hat{\Omega}} = M' \underline{\omega}_m + M' \underline{D} + M' \underline{b}_1 \quad \underline{\hat{\Omega}} = \underline{Y} + \underline{D}' + \underline{b}'_1$$

$$\text{where } M' = (M^T M)^{-1} M^T$$

\underline{Y} : measurement vector in satellite axes

$\underline{D}' = M' \underline{D}$: imaginary drift such that

$$\dot{\underline{D}}' = M' \dot{\underline{D}} = M' \underline{b}_2 = \underline{b}'_2$$

$$E(\underline{b}'_2 \cdot \underline{b}'_2^T) = M' Q_2 M'^T$$

$$\underline{b}'_1 = M' \underline{b}_1 : E(\underline{b}'_1 \cdot \underline{b}'_1^T) = M' Q_1 M'^T$$

The model becomes

p	$= y_1 + d_1 + V_{11}$
q	$= y_2 + d_2 + V_{12}$
r	$= y_3 + d_3 + V_{13}$
\dot{d}_1	$= V_{21}$
\dot{d}_2	$= V_{22}$
\dot{d}_3	$= V_{23}$

Now, for any application, we have to derive the kinematic equation relating the absolute rate vector to the angular rate of the spacecraft

$$\underline{\dot{\Omega}} = M \underline{\dot{q}}$$

Introducing this equation in the previous set leads to the usual state space model

$$\underline{\dot{X}} = A \underline{X} + \underline{Y} + \underline{V}$$

- Low frequency sensors model

The reference sensors are modelled by a discrete process :

$$Z_K = H_K X_K + V_K$$

where

$$\begin{cases} Z & \text{is the sensors output vector} \\ X & \text{is the spacecraft state} \\ V & \text{is a white noise} \end{cases}$$

so that

$$E(V \cdot V^T) = R \delta(t-\tau)$$

- Kinematic filter equations

The linear Kalman filter in its discrete form is processed any time a sensor output is available. This method avoids any matrix inversion; the filter routine is used sequentially and the small delay induced in the processing does not usually lead to significant errors.

Filter equations

$\hat{X}_{k+1/k+1}$	$= \hat{X}_{k+1/k} + K_{k+1} (Z_{k+1} - H \hat{X}_{k+1/k})$
$\hat{X}_{k+1/k}$	$= \phi \hat{X}_{k/k} + G U_k$
K_{k+1}	$= P_{k+1/k} H^T (H P_{k+1/k} H^T + R_{k+1})^{-1}$
$P_{k+1/k+1}$	$= (1 - K_{k+1} H) P_{k+1/k} (1 - K_{k+1} H)^T + K_{k+1} R_{k+1} K_{k+1}^T$
$P_{k+1/k}$	$= \phi P_{k/k} \phi^T + Q_k$

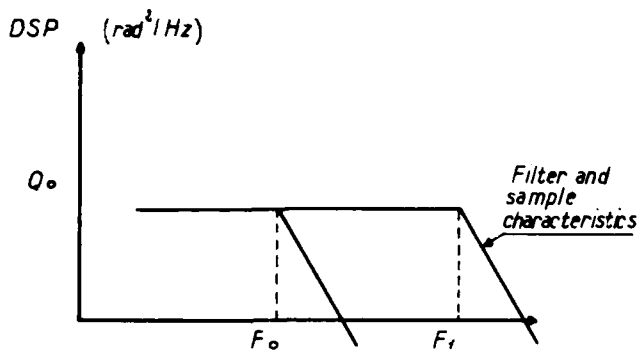


Fig 1 - OPTICAL SENSORS SPECTRUM

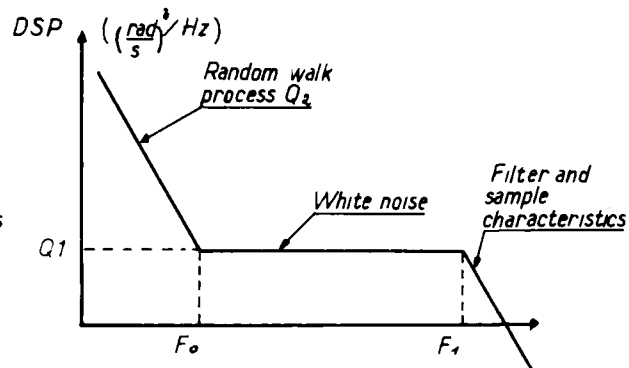
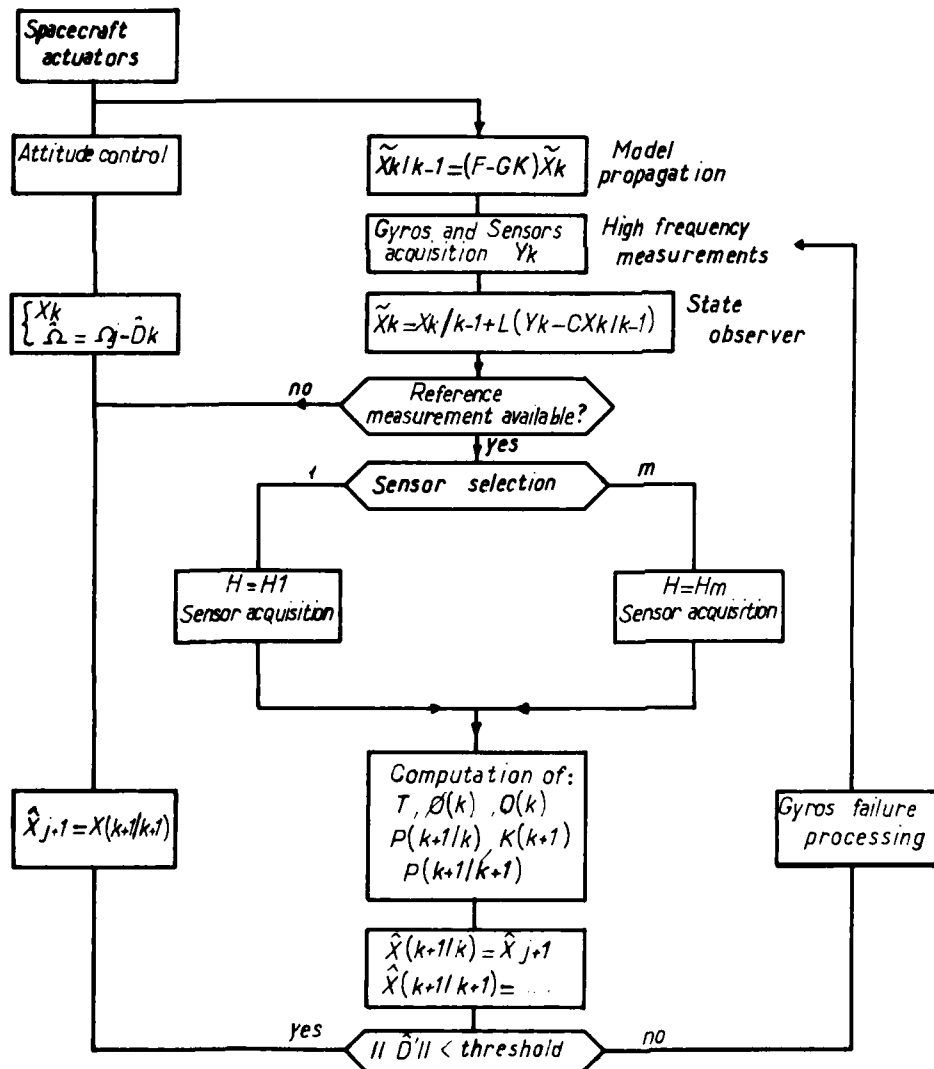


Fig 2 - INERTIAL RATE SENSORS SPECTRUM

Fig 3 - BLOCK DIAGRAM OF THE STATE ESTIMATION PROCESS



2. GENERAL DYNAMIC MODELIZATION:

- τ A dynamic model of the complete satellite is obtained usually by the modal representation, resulting in a linear approximation of the following form :

$$\begin{cases} \dot{\underline{X}} = \underline{A}\underline{X} + \underline{B}(\underline{u} + \underline{D}) \\ \underline{Y} = \underline{C}\underline{X} \end{cases} \quad \begin{cases} \underline{X}^T = (q_1, q_2, q_3, q_4 \dots q_n, \dot{q}_1, \dot{q}_2, \dot{q}_3, \dot{q}_4 \dots \dot{q}_n) \\ \text{with } q_1, q_2, q_3 \text{ as rigid body variables} \\ \underline{u}^T = (u_1 \dots u_m) \\ \underline{u} \text{ is the torque command vector} \\ \underline{D} \text{ is the disturbing torques vector} \\ \underline{Y} \text{ is the sensor output vector} \end{cases}$$

- τ The state estimation filter has the same structure as a Kalman filter (non-minimum order observer). Due to real-time constraints, a reduced dynamic model is used (only modes in or close to the control cut-off frequency are relevant), and constant gains are selected.

The state observer takes the following form :

$$\begin{cases} \hat{\underline{X}}_K = \hat{\underline{X}}_{K/K-1} + L(\underline{Y}_K - \underline{C}\hat{\underline{X}}_{K/K-1}) \\ \hat{\underline{X}}_{K/K-1} = \underline{F}\hat{\underline{X}}_{K-1} + \underline{G}u_{K-1} \\ u_K = -\underline{K}\hat{\underline{X}}_K \end{cases}$$

3. BLOCK DIAGRAM OF THE STATE ESTIMATION PROCESS

The block diagram of the complete state estimation process is depicted in Figure 3, where the inner loop refers to the dynamic filter for short-term control of the spacecraft and the outer loop mainly updates the inertial reference drifts.

ILLUSTRATION OF THE CONCEPT : THREE CASE STUDIES

This concept has been studied through three different research and development programs. As an illustration, both theoretical simulation and physical implementation on the ground hardware have been performed, first to evaluate the long-term behaviour (outer loop performances) and then the short-term behaviour (inner loop performances).

1. LARGE GEOSTATIONARY SPACECRAFT ATTITUDE MEASUREMENT AND CONTROL CONCEPT

Most telecommunication satellites attitude control concepts are based on the conservation of angular momentum; when oriented perpendicularly to the equatorial plane, this property will maintain orientation around the yaw axis (satellite/Earth line). However, solar pressure tends to drift the satellite angular momentum; when an Earth sensor is used to control both pitch and roll axes (perpendicular to the orbit and along the orbit axes), the yaw error is limited in amplitude. This is the orbital gyrocompass effect. Now, when increased power requirements or the need for bigger platforms appear (space platforms, antenna farm satellite, etc.), with the same stringent yaw pointing accuracy, this concept requires an angular momentum unachievable with the current technology.

A direct yaw measurement is then necessary. A number of optical sensors have been investigated (Polaris sensor), but the most straightforward one is the gyroscope.

As an example, in Figures 4 & 5, we describe a configuration where the attitude measurement assembly consists in :

- . an infrared Earth sensor
- . a gyropackage with three components, the sensitive axes of which are skewed in roll/yaw plane
- . a digital Sun sensor, located on X⁺ face of the satellite, providing information around yaw, for a reduced length of time, once per orbit.

Figure 6 depicts the short-term measurement and control concept : the pitch loop is driven by the IRES output, and is decoupled from the other axes. The roll/yaw axes are coupled and driven by gyroscopes rates outputs. The torques commands are a combination of errors and their integral (PI control).

The Kalman filter, taking into account system transition and observation matrices, computes gains to update the new angular errors estimates δ_x and δ_z and the new drift estimates \hat{d}_x and \hat{d}_z .

A typical long-term behaviour is shown in Figure 7, with two periods of Sun observation (30 hrs).

This concept has been physically implemented by MATRA and C.N.E.S. under the IASS contract for INTELSAT. Sensors were mounted on the C.N.E.S. 3-axis controlled air-bearing table, and computation performed on the ground computer (SOLAR 64).

2. EARTH OBSERVATION SATELLITES ATTITUDE MEASUREMENT AND CONTROL CONCEPT

Earth-oriented payload may require both angular rate stability (10^{-4} degree/second) for imagery and a pointing accuracy better than 10^{-2} degree for pattern recognition. The satellite orbit is helio-synchronous so that the Sun vector is at a constant angle from the orbit plane. A variety of configurations have been investigated in preparation for the national and European Earth observation program.

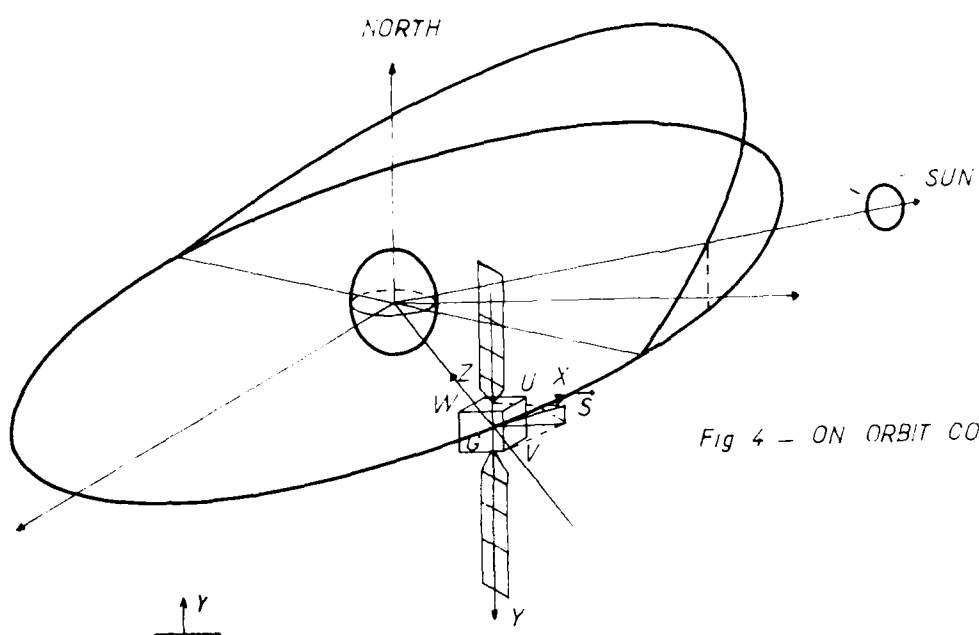


Fig 4 - ON ORBIT CONFIGURATION.

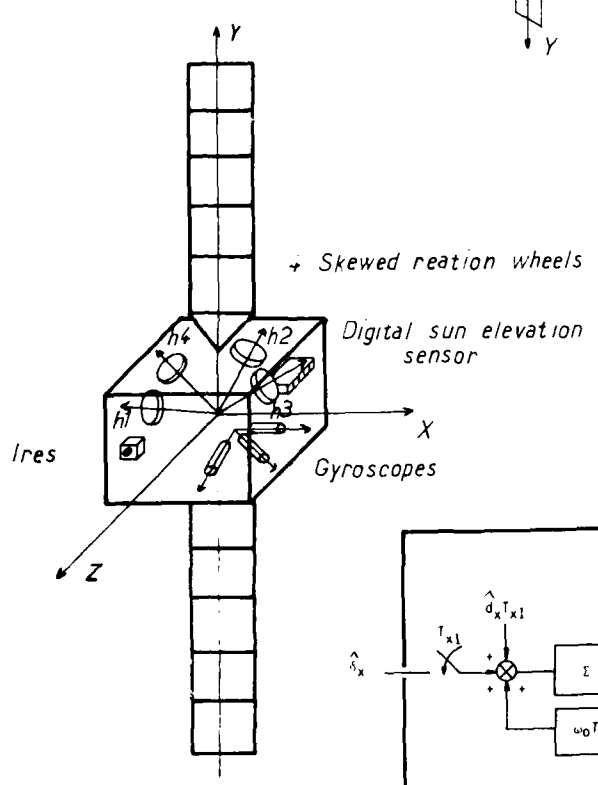
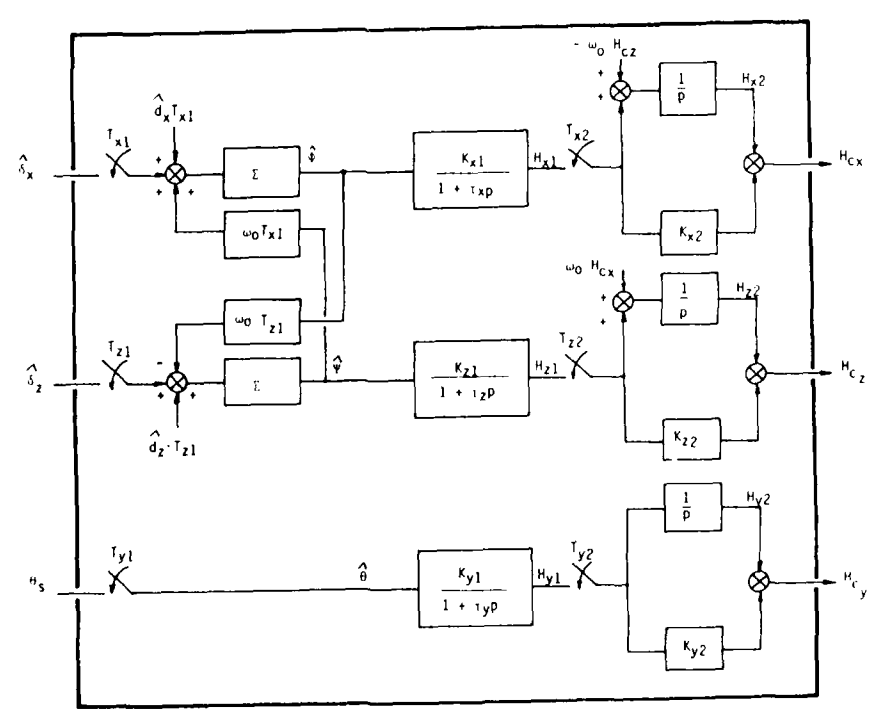


Fig 5 - SENSORS AND ACTUATORS ARRANGEMENT

Fig 6 - NORMAL MODE CONTROL LAWS



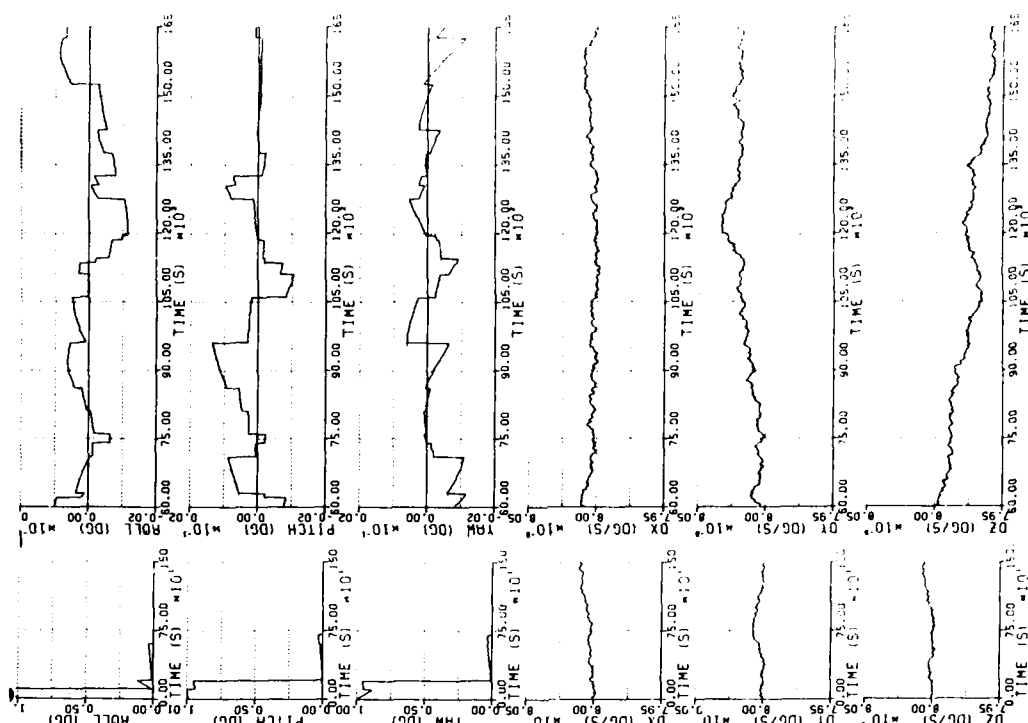


Fig 8 LONG TERM BEHAVIOR FOR AN EARTH OBSERVATION SPACECRAFT

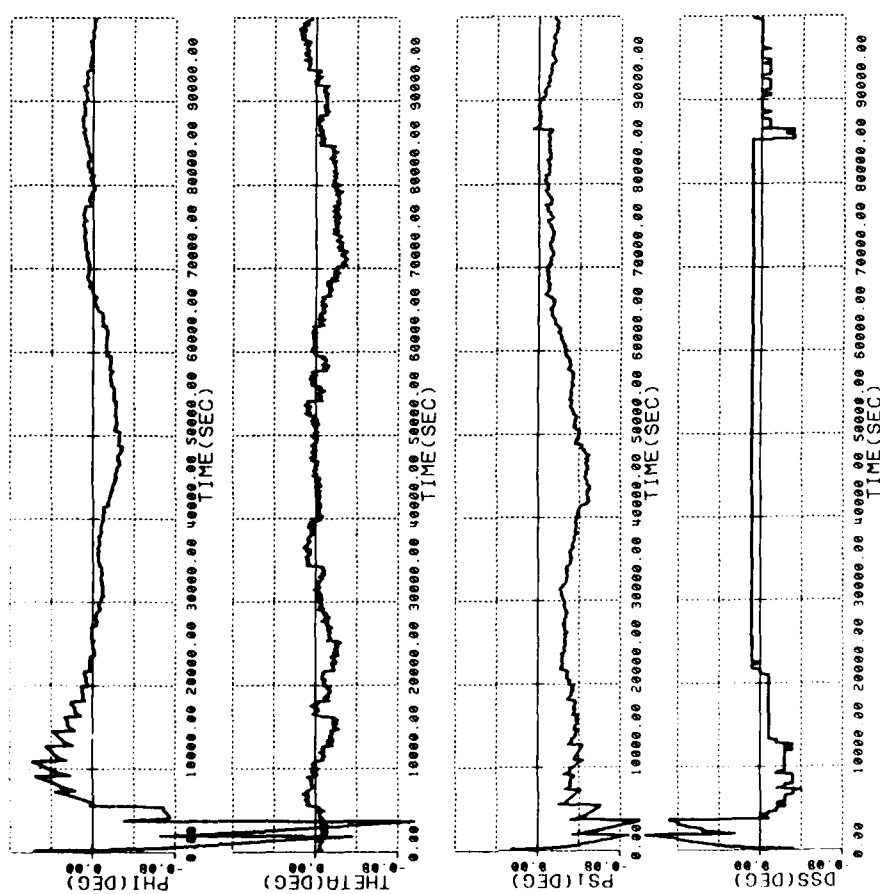


Fig 7 - LONG TERM PERFORMANCES FOR A GEOSTATIONARY TELECOMMUNICATION SPACECRAFT

One of the most promising concepts has been investigated up to the breadboard level; it makes use of the orbital rate which allows the implementation of static CCD linear arrays, scanning the celestial sphere.

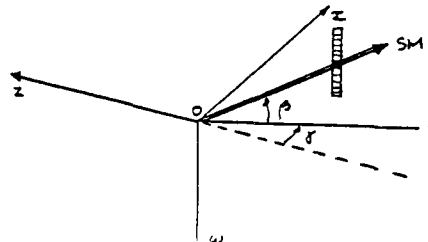
Short-term attitude measurement is performed using a set of six gyroscopes, implemented in a dodecahedral configuration. Control is performed by three magnetic-bearing reaction wheels, oriented along the spacecraft principal axes of inertia.

Measurement equations are described first in some detail, showing the complex star sensors outputs, combining all attitude errors. At one instant, attitude is not fully observable; continuity of the estimation process is ensured using the satellite model prediction. After at least three stars crossings, full observation is possible and complete estimates of spacecraft angular position and drifts are available.

Measurement equations

The star-mapper geometry is defined by angles β and γ .

As the spacecraft orbits the Earth with the Z axis aligned on the geocentric, the star-mappers field of view swaths the celestial sphere. Known stars are observed, and both the instant of crossing and the number of excited CCD elements are recorded.



Error signals are obtained by comparing the sensor output with known star ephemerides :

$$\Delta\chi = \Delta\beta = -\phi \cos \gamma - \psi \sin \gamma$$

$$\Delta t = -\frac{\Delta\gamma}{\omega_0} = \frac{1}{\omega_0} (\phi \tan \beta \sin \gamma + \theta - \psi \tan \beta \cos \gamma)$$

For star-mapper 1 : $\gamma = 0^\circ$; $\beta = 45^\circ$

$$\text{so that : } \begin{cases} \Delta\chi_1 = -\phi + W_1 \\ \Delta t_1 = \frac{\theta}{\omega_0} - \frac{\psi}{\omega_0} + \frac{W_1}{\omega_0} \end{cases}$$

For star-mapper 2 : $\gamma = -90^\circ$; $\beta = -45^\circ$

$$\text{so that : } \begin{cases} \Delta\chi_2 = \psi + W_2 \\ \Delta t_2 = \frac{\phi}{\omega_0} - \frac{\theta}{\omega_0} + \frac{W_2}{\omega_0} \end{cases}$$

A full simulation of spacecraft dynamics, orbit kinematics, star pattern, estimation and control scheme has been performed.

A typical case has been recorded hereunder. Six stars per orbit and per mapper are generated with a non-uniform distribution, in order to have a complex variation of gains and error covariances.

The times of transit stars in SM1 field of view are : 100, 600, 1500, 1800, 2500, 4000 seconds, and for SM2 : 200, 700, 2000, 3000, 5000, 5500 seconds.

Nominal records

Hereunder we recall the nominal data used for reference records :

- τ initial state : $\phi = \theta = \psi = 1 \text{ degree}$
 $d_1 = d_2 = d_3 = 8 \cdot 10^{-5} \text{ deg/sec}$
- τ initial estimation : $\hat{\phi} = \hat{\theta} = \hat{\psi} = \hat{d}_1 = \hat{d}_2 = \hat{d}_3 = 0$
- τ continuous state spectral density : $V_{11} = V_{12} = V_{13} = 10^{-3} (\text{deg/hour})^2/\text{Hz}$
 $V_{21} = V_{22} = V_{23} = 10^{-19} \text{ rd}^2/\text{S}^3$
- τ measurement noise standard deviation
 - . star location : $\sigma = 9 \cdot 10^{-4} \text{ deg}$
 - . transit time : $\sigma = 0.015 \text{ sec}$
- τ initial error covariances
 - . attitude : 1 deg^2
 - . drifts : $0.01 (\text{deg/hour})^2$
- τ orbit altitude : 700 km
- τ number of stars per revolution : 6.

We have recorded the evolution of the state vector : $\dot{x}, \dot{y}, \dot{z}, \dot{d}_1, \dot{d}_2, \dot{d}_3$ in the inertial coordinate values are to be multiplied by the factor indicated to be expressed in the unit in brackets.

A breadboard model of the attitude measurement unit has been built by MATRA under the C.N.E.S. SMAOII contract. A set of four SAGEM gyros has been selected; attitude measurement algorithms were implemented in a digital unit using two CMOS 12-bit microprocessors (HARRIS 6100). The unit was mounted on a C.N.E.S. 3-axis servoed air-bearing table; a ground computer (SOLAR 64) was used to simulate star sensors, the orbital kinematics, and the spacecraft dynamics and control. Nominal performances were verified, and performances below 10 Arc seconds were obtained.

This program has allowed the validation of the concept for a future implementation in space.

3. FLEXIBLE SPACECRAFT STATE ESTIMATION CONCEPT

In a special research and development contract, awarded to MATRA by INTELSAT (IADES contract), we are currently investigating the enhancement of a large communication platform short-term attitude pointing accuracy.

The satellite is described in Figure 9. It consists of a main body, carrying large solar arrays and an offset antenna, gimballed around two axes by an antenna-pointing mechanism and carrying a radio-frequency sensor.

A first investigation of independent control of the main body and large antenna was made, showing the limitations in overall pointing accuracy, mainly due to uncontrolled flexible modes.

When better performances are required, namely in station-keeping, when the solar arrays and antenna flexible modes are excited by the thrusters pulses, a centralized state estimation and control must be performed. Digital processing of gyroscopes rates, the main body IRES and radio-frequency sensor outputs are performed, and modes located in the control bandwidth and slightly above it are observed and controlled.

Figure 10 shows typical simulation results with the enhancement of transient performances.

SATELLITE MISSIONS IMPLEMENTING THE CONCEPT

The optico-inertial concept is currently implemented in a variety of on-going space programs. We now give a short survey of those satellites design.

1. ESA EXOSAT MISSION

ESA has recently launched the EXOSAT spacecraft, the payload of which is devoted to the study of astronomical radiation sources. The attitude measurement assembly is based on a 4-gyroscopes inertial unit (FERRANTI), updated by filtering on the star tracker outputs (SODERN). The spacecraft attitude control is performed by cold gas actuators, and limit cycles of one arc second are currently achieved. Drifts stability has proven excellent (below specifications).

2. SPOT MISSION

The C.N.E.S. Earth observation program first satellite is currently manufactured by MATRA. The attitude measurement sub-assembly is based on a 6-gyroscopes inertial unit (SAGEM, MATRA), an infrared Earth sensor (SODERN), a digital Sun sensor (MATRA). Performances are expected to be lower than 10^{-4} d/sec in angular rate, and 0.15° absolute pointing accuracy.

3. I-SAT MISSION

BAe is now manufacturing the ESA direct TV broadcasting satellite. The attitude measurement assembly consists in a 7-gyroscopes inertial unit (FERRANTI), 5 on the yaw axis for reliability requirements and 2 in the roll/yaw plane. IRES and RF sensors, mounted on APM, are also envisaged. This mission implements both short-term state estimation of flexible modes and long-term updating of the yaw inertial reference.

4. HIPPARCOS MISSION

E.S.A. is now initiating the HIPPARCOS program, aiming at mapping 100 000 stars in our galaxy, the coordinates of which are to be obtained after ground processing at an unachieved accuracy (0.002 Arc second). MATRA is prime contractor for the satellite and the payload. AERITALIA is main contractor for the platform. MATRA and BAe cooperate in attitude measurement and the control subsystem. It consists in a set of 5 gyroscopes, and attitude information is obtained from two slit-type star mappers, located in the payload focal plane. Attitude and drifts updating is made at a random rate (10 to 100 seconds). The performance is expected to be lower than 1.4 Arc seconds on three axes.

A posteriori attitude reconstitution is expected to be better than 0.1 second, a performance achievable owing to the good quality of star mappers and gyroscopes. Attitude control is performed by cold gas actuations.

Several new missions will implement the concept, namely the EURECA platform.

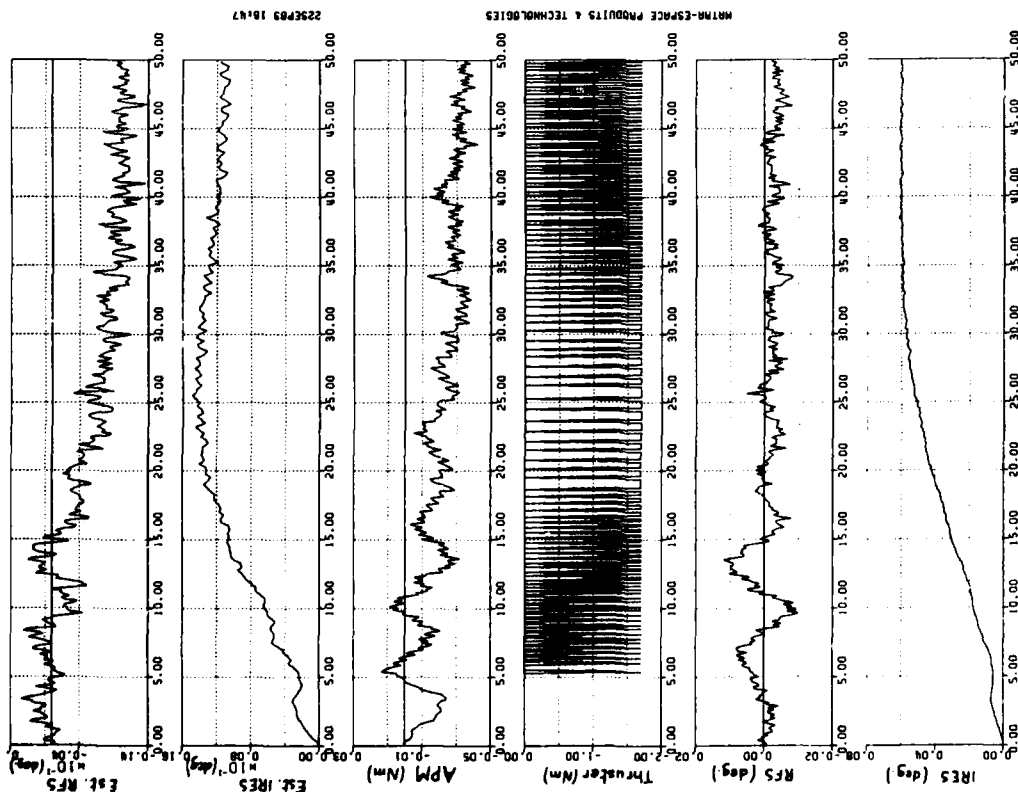


Fig 10 a - PERFORMANCES OF THE CLOSED LOOP SYSTEM

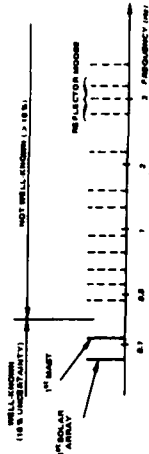
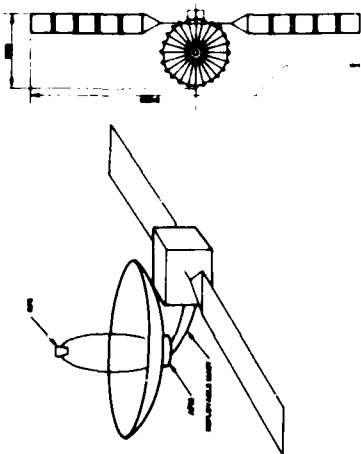


Fig 9 - SATELLITE CONFIGURATION AND MODAL CHARACTERISTICS

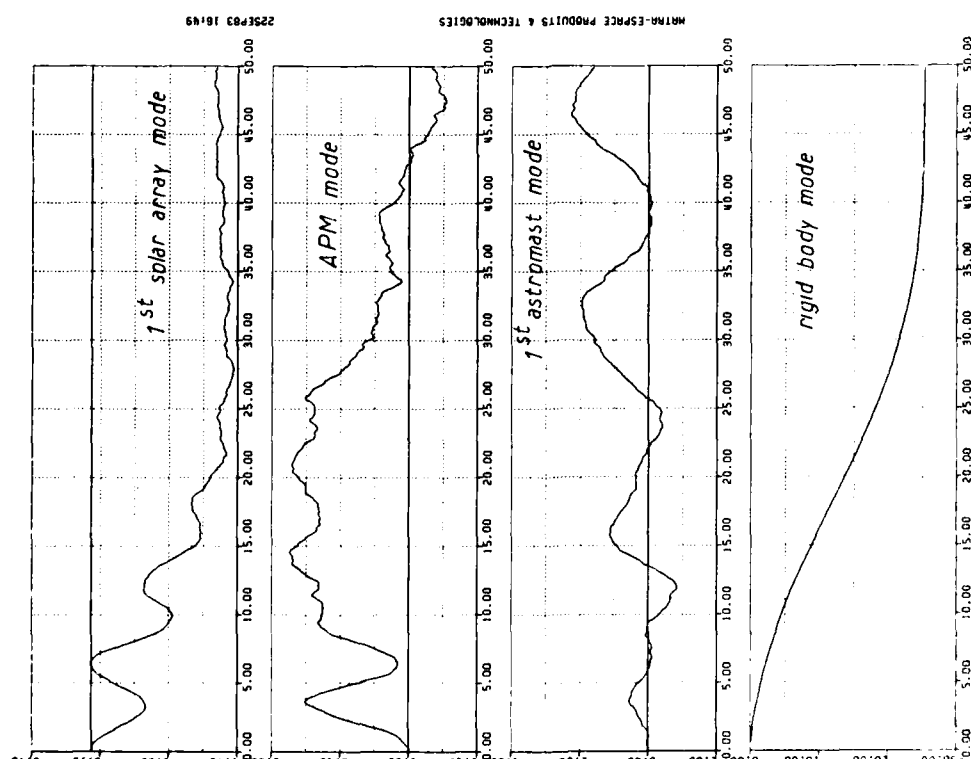


Fig 10 b — ESTIMATED STATES

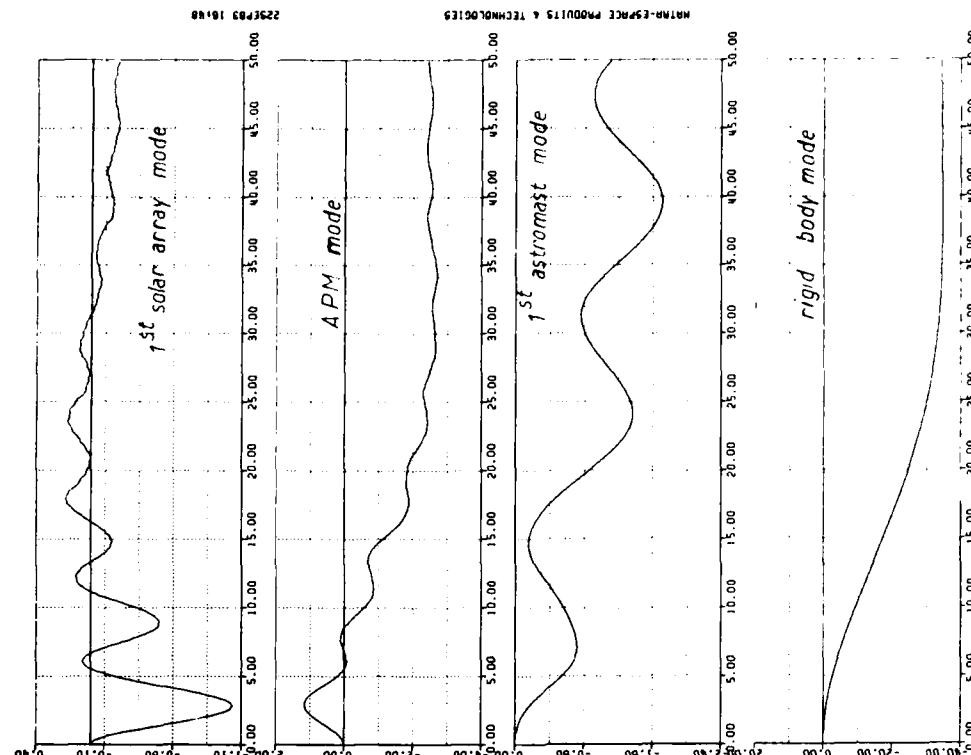


Fig 10 c TRUE STATES

AN EXAMPLE OF PHYSICAL IMPLEMENTATION

In this section, spacecraft on-board data processing is first considered. A description of the "intelligent unit", performing spacecraft state estimation, is given; it is based on standard space qualified components. Newcoming hardware is then considered.

The physical implementation of our concept is made within the frame of the E.S.A standardization of on-board digital processing. Figure 11 shows the selected architecture, centered on a serial data bus (on-board data handling bus), driven by the platform controller, and on which are connected "intelligent" and passive units.

Figure 12 presents the block-diagram of hardware developed by MATRA for C.N.E.S. (SMAO11), with a bi-processor organization. The microprocessor is a HARRIS CMOS 6100, a twelve bits processor, which, until recently, was the only microprocessor available in Europe with radiation hardening sufficient for long-duration space missions.

A first microprocessor is used to manage sensors and to provide the short-term state estimation (gyros projection and integration); the second microprocessor implements the Kalman filter and manages the interface to the external bus. The software size in microprocessor 1 is 3.8 K, including the fixed-point library; with a 5 Hz gyroscopes acquisition frequency, the execution time is 50 ms with a 2 MHz internal clock (safe design). In microprocessor 2, 6.0 K of memory are used, including floating-point library; execution time of the Kalman filter is 3.8 seconds, a delay which has no influence in this application.

Space-qualified hardware is in constant evolution; a new microprocessor is now selected for space application, owing to its resistance against radiations encountered during a 10-year mission at a geostationary altitude. The new units will be fitted with the TEXAS INSTRUMENTS 16 bits SBP 9989.

AUTONOMOUS NAVIGATION

On-board control functions for the recent S/C designs comprise trajectory or orbit automatic determination tasks. These techniques involve the mixing of inertial measurement (accelerometric and gyrometric packages) and external measurements.

In order to simplify this paper, the presentation is restricted to two particular cases which have recently aroused considerable interest :

- . autonomous station-keeping for a geostationary spacecraft
- . autonomous relative trajectory estimation during a low Earth orbit rendez-vous.

1. AUTONOMOUS STATION-KEEPING

In general, autonomous navigation algorithms use two external references (the sun, planets or stars). We will present below a promising method for autonomous orbit determination, which uses only the sun as an external additional measurement, since the Earth reference is almost always available on-board a geostationary telecommunication satellite.

The problem is the determination of the two out-of-plane parameters : longitude of ascending node Ω and inclination i , or the orbit pole vector components :

$$\begin{aligned}\zeta &= i \cos \Omega \quad (\Omega : \text{longitude of ascending node, } i : \text{inclination}) \\ \eta &= i \sin \Omega\end{aligned}$$

Assuming a perfect and circular orbit, the solar declination measurement in a meridian plane of the satellite gives the instantaneous latitude of the satellite (assuming also a perfect control of the pitch axis in the N/S satellite Earth plane)

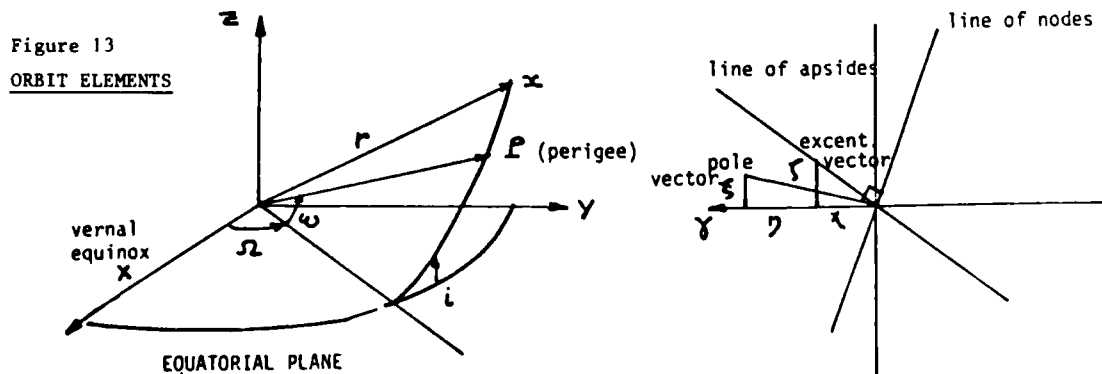
$$l = \zeta \sin \omega_E t + \eta \cos \omega_E t$$

where ω_E is assumed equal to the earth rotation rate and t is known. If we assume also that the node position is known, the orbit inclination can be derived.

Without the assumption on nodes position, two measurements are necessary with different Earth-Sun direction. Thus the determination is only possible over a significant period duration.

Algorithms aspects

In order to show clearly which type of measurement processing is preferred for the orbit determination, a typical approach is developed in a few lines. The orbit determination uses only the Earth-satellite-Sun angle measurement and a reference clock;



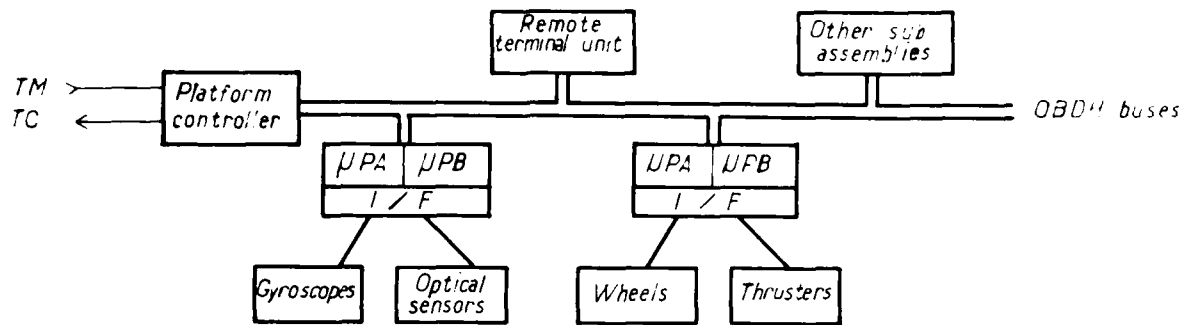


Fig 11 — SCHEMATICS OF STANDARD ON BOARD DATA PROCESSING ARCHITECTURE

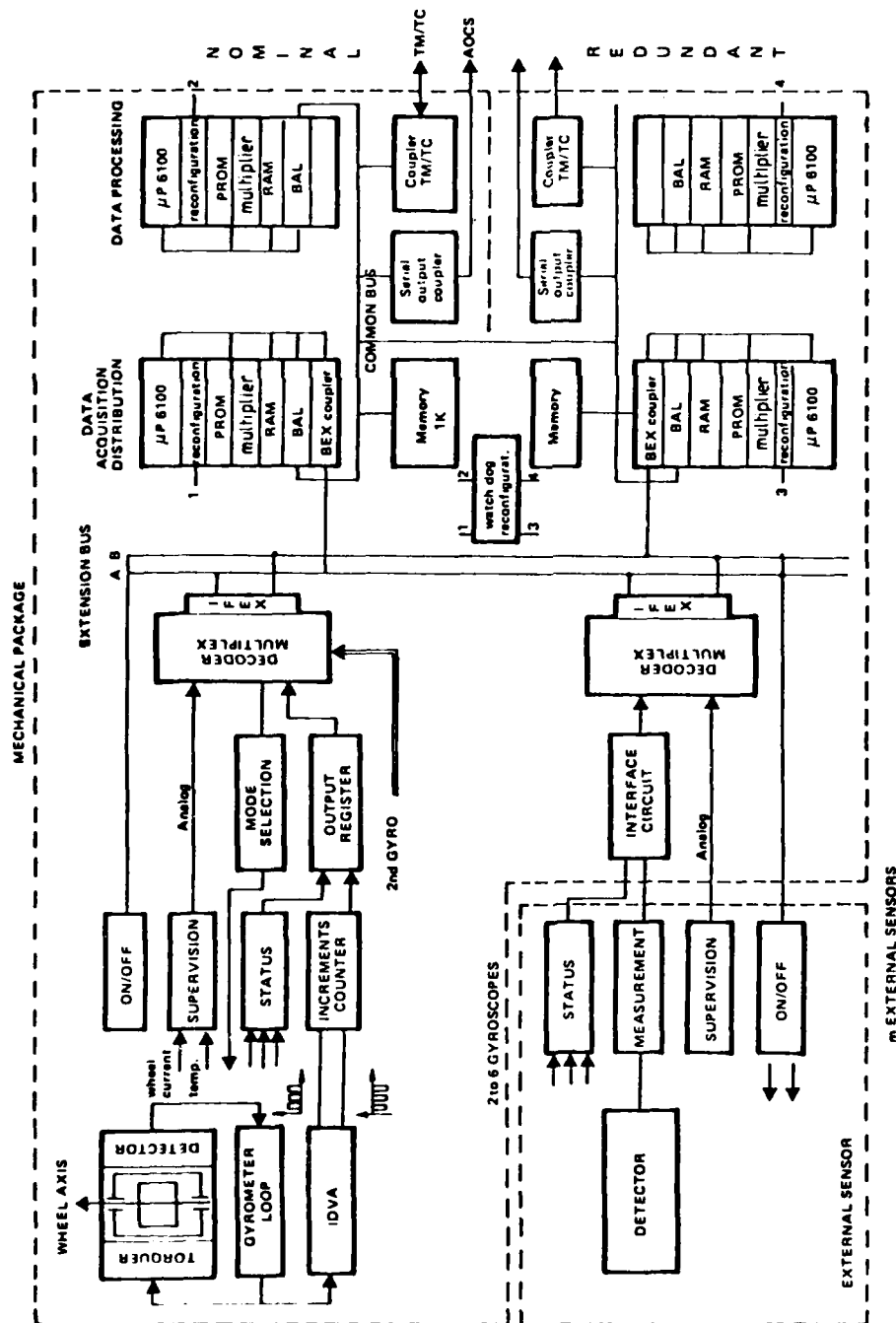


Fig 12 — MODULAR STRAPDOWN OPTICO INERTIAL SYSTEM

Let us take a reference axes system in inertial space, characterized by a North-pointing Z axis and a X axis pointing along the first point of Aries.

The mean longitude of the spacecraft measured relative to the first point of Aries is :

$$\bar{\lambda} = \Omega + \omega + M$$

with

Ω : longitude of ascending node
 ω : perigee argument
 M : mean anomaly.

We also use the pole position vector and the eccentricity vector :

$$\begin{aligned} \eta &= i \sin \Omega \\ \xi &= -i \cos \Omega \end{aligned} \quad \left. \begin{array}{l} \text{pole vector (i : inclination)} \\ \chi = \frac{360}{M} e \cos (\Omega + \omega) \\ \zeta = \frac{-360}{M} e \sin (\Omega + \omega) \end{array} \right\} \quad \begin{array}{l} \text{eccentricity vector (e : eccentricity)} \end{array}$$

The mean longitude is given by : $\bar{\lambda} = \bar{\lambda}_{t=0} + (D_G + \dot{G})t$

where

$\dot{G} = dG/dt$ G : Greenwich sidereal time
 D_G = drift rate relative to Greenwich

Consider the rotation vector W which defined the rotation of the Earth-satellite vector U from nominal to actual one, due to orbit perturbation. The Z component of the vector includes the S/C mean longitude variation (drift rate effect) and the eccentricity effect; the projection of W in the XY plane defines the inclination and the line of nodes of the new orbit :

$$\begin{cases} W_x = -i \cos \Omega \\ W_y = -i \sin \Omega \\ W_z = \Delta\bar{\lambda} + \chi \sin \omega_E t + \zeta \cos \omega_E t \end{cases}$$

with $\Delta\bar{\lambda} = D_G \cdot$

We note S the satellite to Sun vector. The problem consists in the determination of W , knowing only the angle between U and S as seen from the satellite with :

X = Sun-satellite-Earth angle

X_l = Sun-Earth-satellite angle

Angle X_l is given by :

$$X_l = 180^\circ - X - \epsilon$$

with ϵ : parallax correction (up to 0.016°) which can be computed knowing the mean S/C position.

The X angle is computed from the Earth sensor and the Sun sensor outputs. After the parallax correction, the measurement is :

$$y = S^T U \quad (\text{scalar product}).$$

Between the nominal satellite position and the actual one (at a given time t) : $U = U_0 + \tilde{W} U_0$

$$U = (I + \tilde{W}) U_0$$

where W denotes the matrix equivalent of the vectorial product :

$$\tilde{W} = \begin{vmatrix} 0 & -W_z & W_y \\ W_z & 0 & -W_x \\ -W_y & W_x & 0 \end{vmatrix}$$

The measurement is now given by :

$$y = m(t) = S^T U = S^T U_0 + S^T (\tilde{W} U_0)$$

$$y - S^T U_0 = S^T (\tilde{W} U_0) \quad (\text{after perturbation of the mixed product}).$$

In this equation, the left-hand side is a known scalar, while the right-hand is a linear combination of W

$$\Delta y = W \cdot (S U_0)$$

A global determination is based on a Kalman filter which works on the orbit parameters

$$X = (\Delta\bar{\lambda} \ D_G \ \chi \ \zeta \ \eta \ \xi) .$$

The measurement equation is time-variant at the orbit rate :

$$y = \eta(\text{SU})_x + \zeta(\text{SU})_y + [\Delta\lambda + \chi \sin \omega_E t + \zeta \cos \omega_E t] (\text{SU})_z = \text{HX}$$

The evolution matrix ($X = AX$) is :

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \cdot & \cdot & \cdot & \dots \end{bmatrix}$$

The input matrix is given by the main orbit perturbations : Earth gravity harmonics, luni-solar gravity, radiation pressure. For instance, for the radiation pressure in plane effect :

$$\chi = -k_1 S(1 + \sigma)/m \cos \delta_s \sin \alpha_s$$

$$\zeta = -k_2 S(1 + \sigma)/m \cos \delta_s \cos \alpha_s$$

S : effective cross sectional area

σ : average reflectivity

m : spacecraft mass

δ_s/α_s : Sun declination and right ascension.

The orbit correction effect can be also introduced by step on orbit parameters ($\Delta\chi$, $\Delta\zeta$).

The Kalman filter takes the form :

$$\hat{X}_{k/k} = \hat{X}_{k/k-1} + P_{k/k} \cdot H_k^T \cdot R^{-1} (y_k - HX_k)$$

$$\hat{X}_{k/k-1} = \Phi_{k/k-1} \hat{X}_{k-1/k-1} + F_{k-1}$$

(transition matrix)

(input matrix : orbit perturbation, maneuver effect)

$$\left. \begin{aligned} P_{k/k}^{-1} &= P_{k/k-1}^{-1} + H_k^T R^{-1} H_k \\ P_{k+1/k} &= \Phi_{k+1/k} P_{k/k} \Phi_{k+1/k}^T + Q_k \end{aligned} \right\} \text{covariance matrix extrapolation}$$

A deterministic approach will be used rather than a stochastic one : the measurement noise R and the process noise Q covariance matrix are chosen so that the filter has a given response time, for instance

$$Q_k = P_{k/k}/N \quad \text{with } N : \text{ given number of sampling periods.}$$

Sun sensors implementation

High resolution Sun sensors are required, providing a 0.02° accuracy. Digital sensors are candidates:

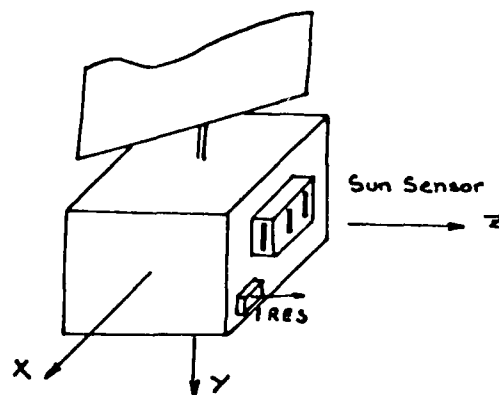
- . two-axis Sun sensors
- . meridian slit-type Sun sensors.

The first class includes new types of CCD Sun sensors. The second class includes digital Sun sensors providing accurate measurements of both Sun elevation angle and the time when the Sun crosses the sensor field of view, which allow to determine the Sun-Earth angle once the relation alignment of Earth and Sun sensor is known.

The number of slits depends on the requirements on the orbit determination duration; for instance, for in-plane parameters, if the parameters determination must be performed in one orbit, 4 slits are probably necessary (4 measurements are necessary for mean longitude, drift rate eccentricity and perigee crossing). However, if we consider that previous information can be used in the determination filter, the number of slits can be decreased. The minimum number of slits is probably 2, if we assume a low variation of parameters (the drift rate is estimated on several orbits and two measurements on the same orbit determine the eccentricity and the perigee crossing). The required FOV for each slit is about 50° .

Figure 14

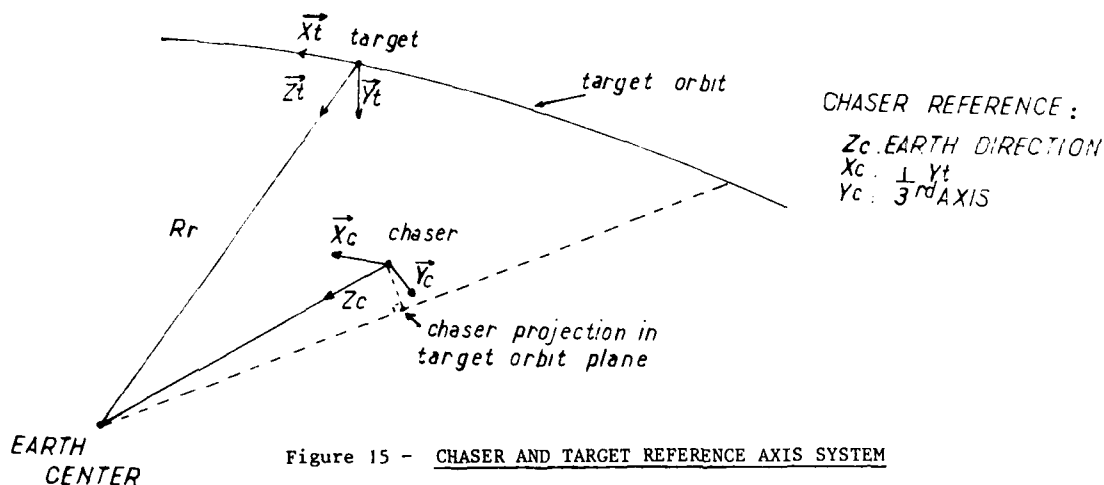
TYPICAL SLIT SUN SENSOR CONFIGURATION



2. AUTONOMOUS TRAJECTORY DETERMINATION DURING RENDEZ-VOUS

The lack of ground assistance (when no TDRSS is available) during a low Earth orbit automated rendez-vous creates an obvious need for on-board navigation. The problem is to identify the relative trajectory w.r.t. the target (assumed here on a circular orbit). A possible approach is presented hereafter.

The chaser spacecraft is Earth-pointed. The rendez-vous sensor is an RF radar with range and angles measurement (monopulse receiver). The problem is to extract the chaser coordinates in the target reference axis, through the RF measurement.



State modelling

A Kalman filter is selected to process radar outputs in order to obtain position estimates. The chaser motion can be written in the form :

$$\dot{\mathbf{X}} = \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{U} + \mathbf{V}$$

$$\mathbf{X} = [x, y, z, \dot{x}, \dot{y}, \dot{z}]^T$$

$$\mathbf{U} = [\gamma_x, \gamma_y, \gamma_z]^T$$

$$\mathbf{V} = \text{perturbation} + \text{model errors.}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_3 & \mathbf{1}_3 \\ 0 & 0 & -2\omega & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\omega^2 & 0 \\ 2\omega & 0 & 0 & 0 & 0 & 3\omega^2 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \mathbf{0}_3 \\ \mathbf{1}_3 \end{bmatrix}$$

Measurement equations

It will be assumed that the radar is boresighted with the X axis of the chaser so that α and ϕ are measured.

Measurement equations are : $m = h(\mathbf{X})$

$$m = \begin{cases} r = (x^2 + y^2 + z^2)^{1/2} \\ \alpha = \text{Arctan} \left(\frac{r_t (r_t - z)}{r_c} - r_t \frac{r'}{x r_t} \right) \\ \phi = \text{Arctan } y \cdot \frac{r_t (r_t - z)}{r_c r'} \cdot \left[x^2 \frac{r_t^2}{r'^2} + \left(\frac{(r_t - z) r_t}{r_c} - r_t \right)^2 \right]^{-1/2} \\ \dot{r} = (x\dot{x} + y\dot{y} + z\dot{z})/r \end{cases}$$

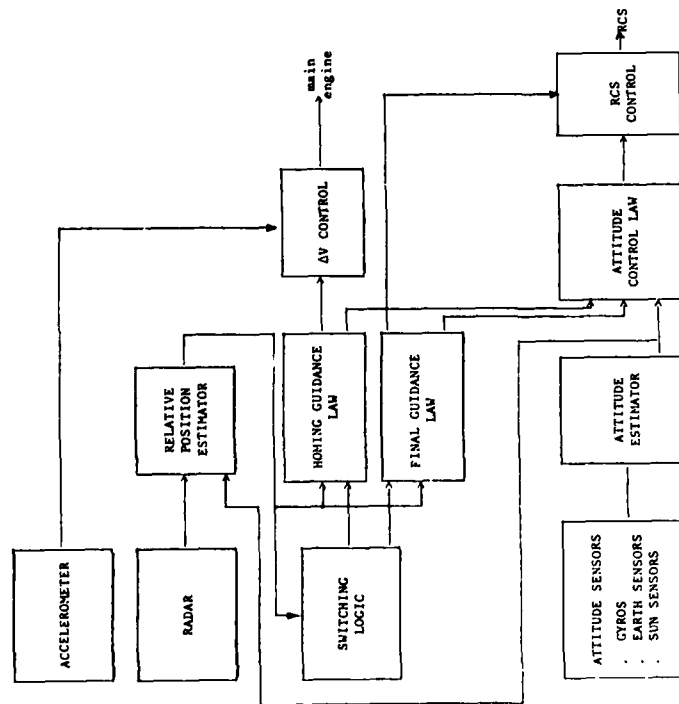
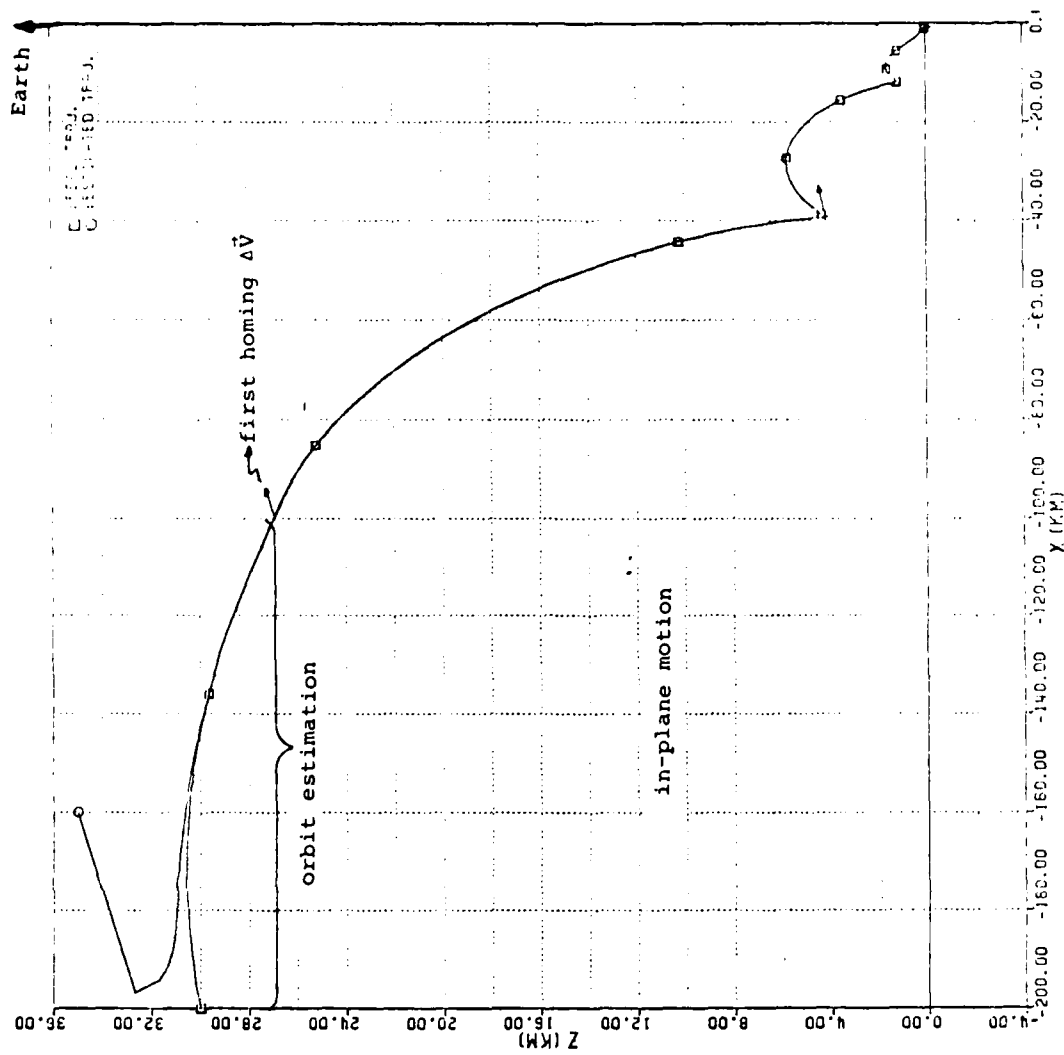


Fig 16 – RENDEZ VOUS TRAJECTORY

The measurement is obviously not linear. The Kalman algorithm (recursive Riccati equation) will then require the Jacobian matrix of the measurement :

$$H_k = \frac{\partial h}{\partial X_k} \quad \left| \quad X_k = \hat{X}_{k/k} + 1 \right.$$

This (4,6) matrix will be computed at each step by a numerical method.

Extended Kalman filter algorithm

The equations are :

$$\begin{cases} \hat{X}_{k/k-1} = F \hat{X}_{k-1/k-1} + G U_{k-1} \\ \hat{Z}_{k/k-1} = h(\hat{X}_{k/k-1}) \\ \hat{X}_{k/k} = \hat{X}_{k/k-1} + K_k (Z_k - \hat{Z}_{k/k-1}) \\ K_k = P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1} \\ P_{k/k} = (I - K_k H_k) P_{k/k-1} \\ P_{k/k-1} = F P_{k-1/k-1} F^T + Q \end{cases}$$

Figure 16 shows simulation results of the estimation process and the resulting rendez-vous trajectory.

CONCLUSION

The optico-inertial concept and its applications has been described in some detail; it has evolved from the initial research and development stage to experience a full development for space.

Extensions are numerous, so that this method is now considered as basic for modern spacecraft navigation and control.

A FILTERED ATTITUDE DETERMINATION SYSTEM FOR SPACECRAFT MEASUREMENT AND CONTROL

by
Maxwell Noton *
Head of Spacecraft Dynamics,
British Aerospace Dynamics Group,
Space and Communications Division,
P.O.Box 5,
Bristol BS12 7QW, England.

SUMMARY

This paper refers to a collaborative project to develop an accurate attitude determination system based on the combined use of gyros and star-crossing detectors. The latter are attractive for earth-pointing satellites but the intermittent nature of the data introduces complications, compared to star tracker systems. Background is given on the hardware, but the emphasis is on formulation of the estimating algorithm (suitable for limited wordlength micro-processors), characterization of gyro drift, and estimation of overall performance both for low-earth and geostationary orbits.

INTRODUCTION

While the requirements of different satellites are a function of the missions, the trend over the last ten years has been to more accurate pointing of spacecraft. Using examples from European missions, ESRO I had pointing to about 10 deg. in 1970, COSB to about 2 deg. in 1975 while the communication satellites of the late seventies and eighties (OTS, ECS and LSAT) have pointing accuracies around 0.1 degrees. However EXOSAT provides a recent example of a scientific satellite with pointing better than 10 arc seconds, and the trend is expected to continue.

The combination of a gyro package and one or more star sensors has become common for accurate attitude measurement systems, but the more accurate systems (in Europe) have been based on a star tracker, i.e. a sensor which retains one or more stars in the field of view for prolonged periods which permits almost continuous resolution of the satellite attitude angles. This paper reports on a UK collaborative project based on the development of a star-crossing detector which is expected to be lighter and consume less power than tracker systems, although it is suitable only for earth-pointing applications or when a satellite has a slow rate of rotation. The method of combining the star crossing data with the inertial rate data from the gyros is however complicated by the intermittent nature of the former, and because the satellite attitude angles cannot be resolved directly from one star transit.

Following an early paper by Jude (ref.9) ESA funded two study contracts, one with Matra and one with RAE, although a proportion of that work was subcontracted to British Aerospace at Bristol (ref.1). As a sequel to these studies, the UK Dept. of Industry (via RAE) initiated the project FADS (Filtered Attitude Determination System) in 1981 with the objective of developing an attitude measurement system, initially to breadboard standards, compatible with the evolving ESA specification for interfacing, viz. MACS or 'Modular Attitude Control System'. The project FADS involves the following collaboration in the UK:

- (1) BAe - Bristol; system studies, microprocessor software and hardware.
- (2) MSDS - Portsmouth; star-sensor, assembly, integration and test.
- (3) Ferranti-Edinburgh; gyro development and testing.

The programme is sub-divided into S and D activities. The former is concerned with the definition of FADS in a more general form suitable for flight applications at a future date, and with extensions to other sensors, etc. The D programme refers to the current specific experimental demonstrations, employing an air-bearing platform originally developed at RAE but now commissioned by MSDS. Although reference is made in this paper to many aspects of the project, it is concerned primarily with system studies and estimation of performance.

* Any views expressed in this paper are those of the author and not necessarily of British Aerospace PLC.

COMPONENTS OF THE SYSTEM

THE STAR SENSOR AND STAR TABLES

The charge-coupled-device (CCD) star sensor was the subject of significant activity at RAE up to 1980. After this date the development was continued by Marconi Space and Defence Systems and is now an essential part of the FADS project. Two such star-crossing detectors are necessary to provide error information accurate about three axes, although one such sensor may be sufficient if a degraded accuracy about one axis is acceptable.

The sensor detects the transit of star images across a linear array of elements lying in the focal plane of the sensor. In an earth-pointing satellite, the star image motion relative to the sensor field of view is due to orbital-rate rotation about the orbit normal as shown in fig.2.1. In one orbit, the field of view of the sensor sweeps out a band of the celestial sphere as shown, of width 9 degrees.

The sensor makes measurements of the timing and position along the array of each star transit, although the basic accuracy is refined considerably by appropriate processing (ref.1, section 4.2). The latter are then compared with the nominal time and position of transit which would be found if satellite attitude errors (and sensor misalignments) were zero, and the discrepancies, Δt and Δx , between the actual and nominal transit data, form the sensor measurements which are related to the satellite attitude errors by the measurement equations below.

$$\left. \begin{aligned} \omega_0 \Delta t &= -\phi \tan \beta \cos \delta + \theta - \psi \tan \beta \sin \delta \\ \Delta x &= \phi \sin \delta - \psi \cos \delta \end{aligned} \right\} \quad (2.1)$$

ω_0 is the satellite orbital angular velocity, and θ , ϕ and ψ are the roll, pitch and yaw angular errors.

The sensor is based on a Fairchild 1728 element CCD array suitable for stars of silicon magnitude 3 or brighter. Future developments may include extended optics, to permit detection of magnitude 4 stars, and even a two-dimensional array when it would become a star tracker providing virtually continuous angular data. RAE (ref.1) estimated that the early version of this star sensor would have an accuracy in both angle and crossing-time of about 6 arc seconds (at 3 sigma level), which is consistent with more recent statements by MSDS, although the accuracy is much better for brighter stars such as magnitude zero. MSDS have indicated that the alignments of the sensors could be calibrated on the ground to within 1 arc second but, due to thermal changes in non-rigid structures, it is generally accepted that in-flight calibrations against ground reference points are desirable.

In order to generate the star tables in a form suitable for the star-crossing detector, it is necessary to carry out the following steps:

- (1) Hold a store of the star ephemeris in heliocentric co-ordinates to an accuracy better than 0.1 arc sec e.g. 204 or 635 stars of magnitude 3 (or brighter), or magnitude 4 (or brighter) respectively.
- (2) Transform co-ordinates for a given star to geocentric co-ordinates and apply corrections (ref.1).
- (3) Transform to orbital co-ordinates, which requires time and the orbital elements (updated from time to time).
- (4) Reject if not in either of the swathes for each sensor, but treating the slit as 4 deg. longer than its actual value e.g. number of stars retained equals 33 (magnitude 3 or brighter) or 42 (magnitude 4 or brighter).

The two options of ground or onboard processing of the above steps has been considered (ref.2). The on-board storage requirements for both options are not excessive, but onboard processing is favoured because it reduces the dependence on the ground link, only occasional updating of orbital parameters being necessary.

The available stars for such a sensor has been determined both in the SIOS and FADS studies (refs. 1 and 2), and an example is summarized in table 2.1. Thus, when working with silicon magnitude 3 stars (or brighter) as with the current star sensor, the typical time between star crossing in a low-earth orbit is 6 minutes and 76 minutes in an equatorial geostationary orbit. However, MSDS have accepted (ref. 17) that, with the low scan rates in a geostationary orbit, the sensitivity would be correspondingly increased to permit detection almost down to magnitude 6 stars, although onboard storage might be prohibitive.

TABLE 2.1 EXAMPLES OF STAR TRANSITS

ORBIT	SILICON MAG.	NO.OF CROSSINGS IN BOTH SENSORS	MLAN TIME BETWEEN TRANSITS, MINS.	WORST CASE, OVER YEAR
98 min., low sun synchronous	3 4	10 - 23 21 - 38	5.9 3.3	S1=2, S2=10 S1=7, S2=11
24 hour equatorial	3 5 - 6	19 160 - 480	76 9 - 3	- -

2.2 RECENT TESTS ON THE FERRANTI 125 GYRO

The drift of the gyro between star sightings is a critical factor determining the overall accuracy of FADS, and therefore emphasis has been placed on obtaining additional data from test runs over periods of 12 hours to 5 days. The Ferranti 125 gyro has of course been employed in a number of space projects such as Ariane, Miranda, Spacelab, IPS, Exosat and Lsat, and detailed information is best obtained from the manufacturer e.g. King's paper (ref.3). A few explanatory remarks may however be convenient for the reader.

The Type 125 is a single-axis floated "rate-integrating" gyro, the float being suspended on jewel and pivot bearings and the wheel in ball bearings. Type 126 has gas bearings but, as a result of space experience and life tests over five years with the Type 125, application of the latter is expected to continue for the foreseeable future. It is designed for use in the rate mode with direct electrical feedback from pick-off to torquer, or in the rate-integrating mode where the feedback from the pick-off measure of spacecraft angle is via the control system and physical rotation. The former is in use in this application with the feedback to torquer adjusted to give a bandwidth of the rebalance loop of 5Hz. This slightly low value was recommended to minimize the danger of a bias arising from high frequency noise being partially rectified by residual nonlinearities.

Extensive tests have been carried out by NLR and Ferranti under an ESA contract (ref.4). The principal series of measurements were with closed loop for different bandwidths of the rebalance loop, the output being a rate signal. The open-loop performance was obtained by a computational conversion, although some measurements were taken. Power spectral densities were calculated from sampled data using the Fast Fourier Transform but the lowest useful frequency was about 2.5mHz or an auto-correlations time constant of 64 secs, i.e. this work gave no information on slow drift. On the other hand it is worth noting the following principal characterization of the high frequency noise:

- (a) Two monochromatic components; amplitude =2.0 arc-sec per sec at 18.6Hz, amplitude =2.0 arc-sec per sec at 27Hz.
- (b) Noise r.m.s. power in band 0-0.05Hz ~0.0031 deg/hr.
and in band 0-5Hz ~0.14 deg/hr.
- N.B. 1 deg/hour = 1 arc-second per sec.
- (c) Integration of the above noise power indicates an r.m.s. angular error in the range 0.02 to 0.05 arc-seconds, the contribution from the monochromatic components being negligible.

The Ferranti 125 gyro will be used in FADS with a sample rate of 10Hz, but the output is actually angular increments over intervals of 1/10th sec. In other words, they are samples of rate averaged over 1/10th sec. intervals, which is roughly like a low-pass filter of bandwidth 4.4Hz (see appendix 8.1). While this may seem to indicate a troublesome time lag for control applications it should be borne in mind, at least for the demonstration with a rigid body load, that wide bandwidth estimates of attitude angles and rates are generated by a subsequent state observer as in Exosat (ref.5).

In the recent 1983 tests (under the UK FADS programme) Ferranti have permitted closer investigation of the frequencies below 50 mHz by recording on tape the rate averaged over sub-intervals of 6 and 60 seconds. This data has been analysed by them to produce additional data on the spectral density, examples of which are fig.2.2 and fig.2.3 for sub-intervals of 6 and 60 seconds respectively. Table 2.2 summarizes the results on four test runs which have also been analysed at BAe in other ways as explained below.

TABLE 2.2 SPECTRAL ANALYSIS BY FERRANTI

RUN NO.	SUBINTERVAL SECS.	MEAN l.f. SPECTRAL DENSITY, (deg/h) ² per Hz	BANDWIDTH TO -3db, mHz
50	6	0.000100	52
60	6	0.000126	
30	6	0.000158	
31	6	0.000158	
40	60	0.00158	5.2
41	60	0.000631	

The last two results indicate greater spectral density at the lower frequencies which is evident in the figures. However the bandwidths of table 2.2 and figs. 2.2 and 2.3 must be remarked upon. It is shown in appendix 8.1 that if a random process of bandwidth ω_0 rad/s is passed through such an averaging filter of subinterval h , then successive averaged pieces of data become independent if $\omega_0 h$ is much greater than unity. The spectral density of the random rectangular waveform is then of the form

$$\left(\frac{\sin \frac{\omega_0 h}{2}}{\frac{\omega_0 h}{2}} \right)^2 \quad (2.2)$$

with a low frequency spectral density equal to that of the original process. The bandwidth of (2.2) to half-power is $0.443/h$ Hz. which for h equal to 6 and 60 secs. yields 7.4 and 0.74 mHz respectively. The spectral densities in figs. 2.2 and 2.3 do not have the exact pattern of (2.2), presumably due to the presence of random components which leave some auto-correlation in successive averaged pieces of data, and also due to Ferranti's additional anti-aliasing filtering. In other words the basic shape of the spectral densities of figs. 2.2 and 2.3 derives from the averaging process etc and is not a property of the original rate data.

The spectral densities of figs. 2.2 and 2.3 exhibit peaks at zero frequency but further investigation below $\frac{1}{2}$ mHz is required. For this BAe have employed additional data provided by Palmstroem of Ferranti Navigation Systems Division. The rate data has been further averaged over intervals of 6, 25.6 and 60 minutes for the three pairs of runs in table 2.2. The data is then in the form of a random time series and is perhaps more convenient for studying properties of slow drift over hours or days. Thus the six series of data are presented in figs. 2.4 to 2.9 and concepts of time-series analysis (after Box and Jenkins, ref.6) are applied to derive an understanding of the properties. At the same time it is remarked that such time series with a subinterval of about 6 minutes are exactly in a form relevant to assessing FADS performance in a low-earth orbit, or about 80 minutes for a geostationary orbit.

The first step is to ensure that one is dealing with a series that approximates to a stationary random process. The data of figs. 2.4 to 2.9 are manifestly not stationary random because (a) evaluations of auto-correlations have spurious large values for well spaced intervals and, (b) the series can be seen to have long-term trends with no meaningful mean value. If the trends were approximately linear we might detrend the data but, because this is doubtful (fig.2.7), the alternative procedure has been adopted of differencing the data, i.e. form a new series $y(n)$ from the original series $z(n)$

$$y(n) = z(n) - z(n-1) \quad (2.3)$$

The new series does not then have large auto-correlations for widely spaced intervals although the theoretical behaviour of certain time-series models is usually only roughly approximated by practical data. Tables 2.3, 2.4 and 2.5 show the first few auto-correlations of the new series $y(n)$, although the numbers are the mean of two sets of data for each of the three intervals. The negative correlation for at least the first spacing interval is evident and therefore second order auto-regressive (AR) and moving average models (MA) have been tried.

The AR model is

$$y(n) = a_1 y(n-1) + a_2 y(n-2) + W(n) \quad (2.4)$$

and the MA model is

$$y(n) = b_1 W(n) + b_2 W(n-1) + b_3 W(n-2) \quad (2.5)$$

The test of a perfect model is that the residuals $W(n)$ are indeed perfectly random, hence the tables show the autocorrelation of the residuals of these models after the coefficients have been fitted (appendix 8.1).

TABLE 2.3 ANALYSIS OF DIFFERENCED DATA FOR 6 MINS.INTERVAL

MODEL	RMS $W(n)$ DEG/HR	AUTO-CORRELATION OF RESIDUALS			
		1	2	3	4
$y(n)$	0.0012	-0.207	-0.283	-0.060	0.060
AR:					
$a_1 = -0.285$	0.0011	-0.078	-0.108	-0.288	-0.084
$a_2 = -0.343$					
MA:					
$b_1 = -0.000395$	0.999	-0.018	0.0002	-0.046	0.064
$b_2 = -0.000468$					
$b_3 = 0.00103$					

TABLE 2.4 ANALYSIS OF DIFFERENCED DATA FOR 25.6 MIN.INTERVAL

MODEL	RMS $W(n)$ DEG/HR	AUTO-CORRELATION OF RESIDUALS			
		1	2	3	4
$y(a)$	0.00071	-0.381	-0.027	-0.124	0.210
AR:					
$a_1 = -0.464$	0.00063	-0.063	-0.113	-0.214	0.137
$a_2 = -0.207$					
MA:					
$b_1 = -0.000020$	0.999	-0.020	-0.033	-0.075	0.179
$b_2 = -0.000312$					
$b_3 = 0.000637$					

TABLE 2.5 ANALYSIS OF DIFFERENCED DATA FOR 60 MIN. INTERVAL

MODEL	RMS $W(n)$ DEG/HR	AUTO-CORRELATION OF RESIDUALS			
		1	2	3	4
$y(n)$	0.0016	-0.275	0.009	-0.154	-0.064
AR:					
$a_1 = -0.310$	0.00148	-0.024	-0.095	-0.204	-0.090
$a_2 = -0.078$					
MA:					
$b_1 = 0.000014$	0.981	-0.023	-0.062	-0.199	-0.095
$b_2 = -0.000455$					
$b_3 = 0.00153$					

The MA model is marginally to be preferred especially for the important 6 min. interval, and has been employed theoretically in section 4 below, although a more convincing demonstration of FADS is considered to be that employing the actual 6 min. time series of drift data in a simulation.

It is interesting in a wider context to speculate if a simple noise and drift model of the gyro in continuous time fits the time-series data when it is converted to the three different intervals employed. First we note a useful model which at least describes the time series with the longer intervals of 25.6 and 60 mins. Define the time series drift data as the sum of two series $z(n)$ and $v(n)$, thus

$$D(u) = z(u) + v(n) \quad (2.6)$$

One series is a random walk, having no mean value,

$$z(n) = z(n-1) + W_1(n) \quad (2.7)$$

and the other series is auto-correlated with zero mean value. Thus

$$v(n) = \beta v(n-1) + W_2(n) \quad (2.8)$$

where $W_1(n)$ and $W_2(n)$ are purely random with zero mean values. If the series $D(n)$ is differenced the residuals are

$$y(n) = D(n) - D(n-1) = W_1(n) - (1-\beta) v(n-1) + W_2(n) \quad (2.9)$$

from which we deduce

$$\frac{y^2}{W_1^2} = \frac{(a+2)\sigma^2}{a\sigma^2}, \quad \frac{W_2^2}{W_1^2} = \sigma^2 \quad (2.10)$$

and the normalised autocorrelations of $y(n)$ are

$$\begin{aligned} \rho_1 &= -(1-\beta)/[(a+2)(1+\beta)] \\ \rho_j &= \beta \rho_{j-1}; \quad j > 1 \end{aligned} \quad (2.11)$$

Compare the autocorrelations of $y(n)$ in tables 2.3, 2.4 and 2.5. The 6 minute data of fig.2.3 does not fit very well due to the large negative correlation ρ_2 but the fit is acceptable to the data for 25.6 and 60 min. intervals using simply $\rho=0$, i.e. a purely random term $v(n)$ added to the random walk $z(n)$. Table 2.6 summarises the parameters of this model.

TABLE 2.6 PARAMETERS OF SIMPLE DRIFT MODELS

INTERVALS MINS.	STAND. DEVIATIONS DEG/HR	
	RANDOM ADDITIVE TERM	INCREMENTS OF RANDOM WALK
25.6	0.000438	0.000346
60	0.000839	0.00107

The above models might be useful for simulation purposes but they are not mutually consistent in terms of an underlying continuous time model. For example the implied low-frequency spectral density is not the same and the increments of the random walk are not in the right ratio for the two intervals. One complication is doubtless that averaging introduces some serial correlation in a random walk model. No further efforts have therefore been made to fit the data to a single continuous time model, but the presence of a significant drift component that behaves as a random walk is clearly an important feature to bear in mind.

2.3 MICROPROCESSOR DEVELOPMENTS

In summarising hardware and software developments at British Aerospace (Bristol), it is important to distinguish between the S (System) and D (Demonstration) programme of FADS. The former is more general and refers to a future system for flight application, whereas the D programme is concerned only with the specific system and hardware for the demonstration on the air-bearing platform already mentioned in the introduction.

Thus the hardware for the S Model (future flight version) cannot yet be defined, although for the next few years a likely candidate is the 16-bit Texas 9989 microprocessor. It has the special Texas so-called "i²L" technology to provide an acceptable degree of resistance to space radiation and has recently been adopted for LSAT. At present, flight software has to be written in assembler code but ESA may adopt a suitable high level language at a future date e.g. a recent invitation from the agency to tender for studies of possible high level languages.

The D model hardware is on the other hand merely to "elegant breadboard" standards. Early comparative studies restricted the choice to the following microprocessors: Texas 9989, Intel 8086, Zilog Z8000 and the Motorola 68000. The choice was not obvious but in fact the Intel 8086 was selected. It has already been classified by ESA as "Category 3, qualifiable" and claims have been made (Ref.7) that it is suitable for low orbit missions of limited life-times. Furthermore, it is understood that a CMOS (radiation tolerant) version of the Intel 8086 processor is available.

As regards software for the D programme, Pascal has been chosen for the following reasons:

- (a) it is processor independent but available for all the possible processors for FADS.
- (b) its structure is such that it can be employed to define programs at various levels of abstraction without ambiguity.
- (c) it has been defined as an international standard, hence transportability of programs.

The current developments at BAe-Bristol on the FADS hardware and software are now summarised. First, following analysis and simulations described in sections 3 and 4, the application software is developed and tested (in Pascal) off-line. This includes star table generation and updating, the estimation algorithm, and the state observer. Processing of signals from the star sensor to enhance accuracy is carried out in the MSDS electronics of the star sensor. Similarly the Ferranti 125 gyro has its own built-in electronics to provide the rate at 1/10th sec. intervals (averaged throughout such intervals), although fine correction for drift rate is implemented inside the FADS processor. The final editing of the application software is carried out on the Intel Series III Microcomputer Development System, before compilation and transfer to an Intel Single-Board-Computer (SBC) 86/05.

The operating system of the Series III would of course be far too elaborate for the SBC and therefore a much simpler operating system has been developed. It is based upon an ESTEC design documented in 1979(ref.8). Interfacing with the gyro unit, star sensor and actuators is via a serial digital data bus designed according to the new ESA specifications known as MACS (modular attitude control system), or at least a provisional version of the specification. The FADS demonstration will become therefore a candidate for subsequent testing and demonstration of MACS.

Before the FADS processor is delivered as an SBC 86/05, for use by Marconi Space and Defence Systems on the experimental set-up, it will be tested by real-time simulation at Bristol. For this, test software is being developed to simulate (in real-time) the star sensor, the gyro units and the actuators and dynamics of an illustrative configuration e.g. LSAT in the normal mode. This test software will also be implemented on the Intel single-board-computers. However, whereas implementation of FADS is being deliberately restricted to the basic Intel 8086 which has no floating point hardware operations, the test simulation boards can take advantage of the greater numerical processing capability of the extra 8087 board, which does include floating point operations in hardware. Control of the simulation experiments and data logging will be by means of a Texas 990 minicomputer.

FORMULATION OF THE FADS ALGORITHM

DISCUSSION

The intermittent nature of the data from a star-crossing detector makes the process of combining gyro and star sensor measurements much more involved than when continuous data is available e.g. from a star tracker or RF sensor. For this reason the application of the Kalman Filter has been common (ref.9) and was employed in the earlier SIOS (strapdown inertial optical system) studies by RAE and BAe-Bristol (ref.1). A sixth-order state estimator was employed for three satellite attitude (Euler) angles and three drift rates of the gyros. Note that, because the star crossings occur intermittently, the non-stationary form of the Kalman Filter must be employed and the covariance matrix cannot be pre-computed. This, together with the inclusion of the frequently sampled gyro measurements in the propagation of the state vector, made the onboard processing load rather heavy.

In this project (the sequel to SIOS) the estimation problem has been reformulated to take advantage of the fact that the non-drift random component of gyro error, when expressed as an angle, is very much less than the star sensor error, viz about 0.03 arc second as noted in section 2.2 above. The problem can then be expressed as one of estimating and updating six slowly varying near-constant parameters.

Propagation forward of the filter state equations does not then arise, although a matrix analogous to the covariance matrix must still be propagated forward at each star sighting.

The planning horizons of this project have been shortened to the extent that it was considered necessary that implementation should be possible on the kind of 16-bit microprocessor that is already space qualified for long lifetime high orbits. Examples are the Ferranti 100L and the Texas 9989 for which only hardware fixed-point multiplication could be assumed. Further changes to the algorithm have therefore been formulated to make it less sensitive to limited wordlength fixed-point operations. The final result, as detailed below, is recursive weighted least-squares but in a "square root" form, similar to an early algorithm of Potter (ref.10) in which no matrix inversion is required. The final form bears naturally a marked resemblance to the Kalman Filter in a factorised form; see for example Kaminski's survey paper (ref.11). A ninth order form of estimator was studied at an earlier stage (ref.2) but was abandoned when the more elementary processor was assumed. It included estimation of three out of four sensor misalignments, but all the estimations were noisier and it was considered of doubtful benefit. Calibration of misalignments with reference to ground reference points is henceforth assumed.

Perhaps reference should also be made to publications concerned with quaternions (4 Euler parameters) instead of the 3 Euler angles. They are usually in the context of either strapdown inertial navigation systems (ref.13) or attitude estimation systems associated with a three-gimbal inertial platform (ref. 14,15), also forming part of a guidance system and processing almost continuous angular data from a star tracker. In either case the use of quaternions is shown to have advantages in carrying out frequent transformations of vectors from one system to another, because trigonometric functions are obviated. It is a very different application of Kalman Filtering, compared to the simpler problem here, for which the use of quaternions is not considered to be attractive in view of the extra complications of a singular seven-dimensional covariance matrix (ref.15).

3.2

SUMMARY OF ANALYTICAL DEVELOPMENT

In order to concentrate on the principal features small error terms are omitted below, although the results of error analyses (ref.2) are quoted in section 4 on performance. The vector V_g of gyro output rate signals is (ref.1) for unit scale factor

$$V_g = \omega - d + \begin{bmatrix} 0 \\ \omega_0 \\ 0 \end{bmatrix} \quad (3.1)$$

where ω is the vector of true inertial satellite rates and d is the vector of drift rates. The third term is a torquer signal to offset the orbital rate.

Let $\alpha^T = (\phi, \theta, \psi)$ be the satellite Euler angles in roll, pitch and yaw relative to the orbital earth-pointing rotating axes, then

$$\dot{\alpha} = \omega + \begin{bmatrix} \omega_0 \psi \\ \omega_0 \\ -\omega_0 \phi \end{bmatrix} \quad (3.2)$$

Eliminating ω ,

$$\dot{\alpha} = P\alpha + V_g + d \quad (3.3)$$

where

$$P = \begin{bmatrix} 0 & 0 & \omega_0 \\ 0 & 0 & 0 \\ -\omega_0 & 0 & 0 \end{bmatrix} \quad (3.4)$$

The solution of (3.3) is

$$\alpha(t) = \Phi(t)\alpha_0 + \int_0^t \Phi(t-\tau)[V_g(\tau) + d(\tau)]d\tau \quad (3.5)$$

where

$$\Phi(t) = \begin{bmatrix} \cos \omega_0 t & 0 & \sin \omega_0 t \\ 0 & 1 & 0 \\ -\sin \omega_0 t & 0 & \cos \omega_0 t \end{bmatrix} \quad (3.6)$$

Note that (3.3) is often approximated by omitting P , which is sufficiently accurate if feedback control is applied to zero the satellite angles, but the exact form is retained here to permit application to satellites without tight attitude control.

Put

$$v(t_n) = \int_0^{t_n} \Phi(t_n - \tau) V_g(\tau) d\tau \quad (3.7)$$

and approximate $d(\tau)$ in (3.5) by a constant drift rate d . Equation (3.5) can be written

$$\alpha(t_n) = \Phi(t_n) \alpha_0 + \Phi_i(t_n) d + v(t_n) \quad (3.8)$$

where

$$\Phi_i(t) = 1/\omega_0 \begin{bmatrix} \sin \omega_0 t & 0 & (1 - \cos \omega_0 t) \\ 0 & \omega_0 t & 0 \\ -(1 - \cos \omega_0 t) & 0 & \sin \omega_0 t \end{bmatrix} \quad (3.9)$$

Put

$$x^T = (\phi_0, \theta_0, \psi_0, d_1/\omega_0, d_2/\omega_0, d_3/\omega_0) \quad (3.10)$$

and in (2.1)

$$(y_n, z_n) = (\omega_0 \Delta t_n, \Delta x) \quad (3.11)$$

Then, in the absence of measurement errors, the results can be condensed to

$$\left. \begin{aligned} y_n' &= y_n - h_n^T v(t_n) = b_n^T x \\ z_n' &= z_n - k_n^T v(t_n) = c_n^T x \end{aligned} \right\} \quad (3.12)$$

where h_n, k_n, b_n, c_n are quoted in appendix 8.2.

In equations (3.12), note that y_n, z_n are the star-sensor errors subject to typical errors of several arc-seconds, whereas the associated terms in $v(t_n)$ arise from integrations of the gyro rate signals between star sightings. Random errors in the latter are the order of 0.03 arc-second i.e. negligible compared to star-sensor errors. Therefore these terms are merely to produce the motion connected measurements (y_n', z_n') . Given a number of such star-sensor measurements, equation (3.12) can form the basis of a weighted least-square fitting procedure to estimate the vector x at (3.10), hence permitting calculation of drift rates, satellite angles by (3.8) and rates by (3.3).

The development of a sequential solution to this problem of weighted least-squares fitting is noted in appendix 8.2. The weighting of recent data is introduced to place decreasing emphasis on old data and hence an estimator that permits updating for parameters that change slowly. Thus, with weighting decreasing as a geometric progression of ratio μ , the time constant of the 'memory' is

$$-1/\log_e \mu \quad \text{sightings} \quad (3.13)$$

From appendix (8.2) the sequential solution for the estimates is

$$\hat{x}_n = \hat{x}_{n-1} + B_n [b_n (y_n - b_n^T \hat{x}_{n-1}) + c_n (z_n - c_n^T \hat{x}_{n-1})] \quad (3.14)$$

where

$$B_n = \frac{1}{\mu} B_{n-1} - \frac{1}{\mu} B_{n-1} H_n [H_n^T B_{n-1} H_n + \mu I]^{-1} H_n^T B_{n-1} \quad (3.15)$$

and

$$H_n = [b_n \ c_n] \quad (3.16)$$

Thus, the matrix inversion in (3.15) is merely 2×2 .

The above procedure generates mean-square consistent estimates subject to certain formal restrictions (ref.12), regardless of the initial conditions imposed upon the sequences \hat{x}_n and B_n . Furthermore it can be shown (ref.12) that the use of initial conditions

$$\hat{x}_0 = 0, \quad B_0 = R_0/\sigma^2 \quad (3.17)$$

generates the conditional expectation of \hat{x}_n given the measurements, in the case of independent normal measurement errors of variance σ^2 , \hat{x}_0 being on a priori normal distribution with zero mean and covariance R_0 . Equation (3.17) is therefore a useful guide for starting the above sequential estimator.

3.3 SQUARE ROOT FORMULATION

The matrix B_n in equation (3.15) is proportional to the a posteriori covariance matrix of errors in estimating the six components of x , at (3.10). For very accurate estimates this matrix approaches singularity, hence the possibility of sensitivity to limited word length, especially when rapid convergence occurs. The symmetric matrix B_n should remain positive definite but this cannot be assumed in demanding situations with limited word lengths. Therefore decompose B_{n-1} as

$$B_{n-1} = S_n S_n^T \quad (3.18)$$

where the non-unique matrix S_n is referred to as the 'square root matrix'. It requires approximately only half the accuracy of that required to store B_{n-1} and, whatever S_n it is evident that B_{n-1} is always at least positive semi-definite. An example of (3.18) is the Cholesky decomposition in which S_n is upper right triangular, but no such restriction is implied above.

The transformation of algorithm (3.14, 3.15) to the square root form is derived in appendix 8.2 and the result can be stated as follows.

$$\hat{x}_n = \hat{x}_{n-1} + \bar{a}_n^{-1} S_n e_n [y_n - b_n^T \hat{x}_{n-1}] \quad (3.19)$$

where

$$\left. \begin{aligned} a_n &= e_n^T e_n + \mu \\ e_n &= S_n^T b_n \end{aligned} \right\} \quad (3.20)$$

and

$$S_{n+1} = S_n [I/\sqrt{\mu} - e_n e_n^T / \alpha_n] \quad (3.21)$$

$$\alpha_n = \sqrt{\mu} [\mu + e_n^T e_n] + \mu \sqrt{\mu + e_n^T e_n} \quad (3.22)$$

The above solution is written out only for scalar measurements, i.e. it implies that the pair of measurements at (3.12) are treated as two separate scalar measurements. The actual fitting problem involves processing the simultaneous occurrence of two measurements (star crossing angle and time). If the data is not weighted ($\mu = 1$) then there is no objection to treating a pair of measurements as two successive scalar measurements. The above algorithm has been employed even for values less than unity, because anticipated values for μ are close to unity, when the slightly different weighting for measurements at the same time would be a negligible consideration. In this way the square root formulation can be applied in its simplest form, with no matrix inversion.

PERFORMANCE

SOURCES OF ERROR

The relevant sources of error are summarised below.

- (a) Star sensor and associated errors (one sigma):

Star sensor 2.0 arc sec (MSDS provisional)

Star e phemeris 1.0 arc sec (Section 4.3, ref.1)

Orbit determination 1.4 arc sec (Appendix M, ref.1)

Composite rms error, = 2.6 arc sec (4.1)

Sensor misalignments assumed to be calibrated with reference to ground reference points.

- (b) Gyro random noise, above 2mHz:

Order of 0.03 deg (see section 2.2)

- (c) Gyro drift, below 2mHz:

See section 2.2

- (d) Gyro misalignment:

Better than 2 arc min (ref.16)

- (e) Gyro scale factor:
0.1 per cent (ref.16)

Effects (b), (d) and (e) will be very small (ref.2) and are henceforth ignored. Instead attention is concentrated below on the principal sources of error, viz. the random composite star-sensor error and gyro drift. Note that errors that might arise through using a state observer are not considered here. As mentioned in section 2.3, an observer is included in the D model for the experimental demonstration with a rigid body load, but significant difficulties would be encountered (due to parameter mismatch) if this were attempted for a flexible load.

4.2

EFFECT OF COMPOSITE STAR-SENSOR ERROR

The covariance matrix of errors in estimating the six components of \mathbf{b} in (3.10) is theoretically equal to B_0/σ where σ is the r.m.s. composite star-sensor error. Table 4.1 shows therefore the standard deviations of the errors in estimating angles calculated in this way for the illustrative sensor geometry (fig.2.1)

$$\begin{aligned}\beta_1 &= 20 \text{ deg}, \gamma_1 = 0 \text{ deg.} \\ \beta_2 &= -20 \text{ deg}, \gamma_2 = 82.4 \text{ deg.}\end{aligned}$$

These sensors are orthogonal and it is assumed that this desirable orientation would be achieved at least approximately. Table 4.1 was obtained by applying the equations of the square-root sequential estimator of section 3.3. The results can however be interpreted only approximately because they depend on the sequence of sensors (1 and 2) and the distribution of crossings over the length of the slit. Memory time is calculated according to (3.13). Notice that the standard deviations are quoted in a dimensionless form, hence σ at (4.1) must be introduced.

The $\bar{\alpha}$ and $\bar{\Delta d}$ of table 4.1 are defined as follows:

$$\begin{aligned}\bar{\alpha} &= \sqrt{(\hat{\phi}^2 + \hat{\theta}^2 + \hat{\psi}^2)/3} \\ \bar{\Delta d} &= \sqrt{[(\Delta \hat{d}_1/\omega_0)^2 + (\Delta \hat{d}_2/\omega_0)^2 + (\Delta \hat{d}_3/\omega_0)^2]}\end{aligned}\quad (4.2)$$

where ω_0 is the orbital rate.

TABLE 4.1 ILLUSTRATIVE ERRORS DUE TO STAR SENSORS

SENSORS	WEIGHTING	MEMORY, NO. SIGHTINGS	STAND. DEVIATIONS			Eqn (4.2)	
	μ		$\hat{\phi}/\sigma$	$\hat{\theta}/\sigma$	$\hat{\psi}/\sigma$	$\bar{\alpha}/\sigma$	$\bar{\Delta d}/\sigma$
2	0.95	19.5	0.47	0.67	0.47	0.545	0.270
2	0.90	9.49	0.63	1.29	0.64	0.907	0.360
2	0.80	4.48	0.98	3.69	0.96	2.27	0.522
2	0.70	2.80	1.39	7.74	1.29	4.60	0.699
2	0.50	1.44	2.85	24.4	2.26	14.2	0.854
1	0.95	19.5	5.09	2.01	0.42	3.17	2.94
1	0.90	9.49	6.53	2.71	0.57	4.10	3.78

Because star crossings occur typically over the length of the slit the vectors b_0 and c_0 of (3.12) vary from one crossing to the next. Estimation is consequently quite feasible with only one star sensor, although the accuracy about an axis through the centre of the slit is naturally degraded. Table 4.1 includes two cases rerun with only sensor 1 operative. Note that the results of table 4.1 are independent (as quoted) of the kind of orbit, such as low-earth or geostationary, because the memory factor is expressed as number of sightings independent of the time between star transits, and $\bar{\Delta d}$ of equation (4.2) is dimensionless.

Following equation (3.8) the estimated angle is:

$$\hat{\alpha}(t) = \Phi(t)\hat{\alpha}_{o,n} + \Phi_i(t)\hat{d}_n + v(t) \quad ; \quad t_n \leq t < t_{n+1} \quad (4.3)$$

where the actual drift rate between t_{n-1} and t_n is d_{n-1} , and the estimate \hat{d}_n is obtained at t_n . Star sensor errors give rise to fluctuations in the estimates $\hat{\alpha}_{o,n}$ and \hat{d}_n , as examined in Section 3, but we are concerned here with the effect of gyro drift. It gives rise to additional errors in the estimates and (in between star transits) an error due to imperfect correction for the gyro drift rate. The estimated angular rates are incidentally

$$\hat{\dot{\alpha}}(t) = P\hat{\alpha}(t) + v_g(t) + \hat{d}_n \quad ; \quad t_n \leq t < t_{n+1} \quad (4.4)$$

See Figure 4.1 for a block schematic of the whole FADS estimator.

Compared to (4.3), the true vector of Euler angles is given by (3.5) and, hence, by subtraction the error is

$$\Delta\hat{\alpha}(t) = \Phi(t)[\hat{\alpha}_{o,n} - \alpha_o] + \Phi_i(t)\hat{d}_n - \int_0^t \Phi(t-\tau)d(\tau)d\tau \quad (4.5)$$

When the interval between sightings is much smaller than the orbital period, $d(\tau)$ in (4.5) can be approximated by the mean value in each subinterval. This permits the following expression which is written out for the worst case of $t = t_{n+1}$, i.e. just before a fresh star transit.

$$\Delta\hat{\alpha}(t_{n+1}) = \Phi(t_{n+1})[\hat{\alpha}_{o,n} - \alpha_o] + \Phi_i(t_{n+1})\hat{d}_n - p(n+1) \quad (4.6)$$

Equation (4.6) is convenient for employing empirical gyro drive data (as provided by Ferranti) averaged over a typical subinterval, such as 6 minutes for a low-Earth orbit. Thus, by inserting such sequences into an off-line simulation of the FADS square-root sequential estimator of Section 3, it has been possible to compute the resultant errors due to gyro drift fluctuations. The results depend upon the geometry and sequence of transits which would not naturally be at equal sub-intervals.

The following Table 4.2 can, therefore, be regarded only as illustrative. The rms values were calculated over approximately 100 samples, after allowing the first sixteen samples for convergence of the estimator. In fact, provided large diagonal elements are used to prime the matrix S of the square-root filter, rapid convergence was achieved in only a few transits more than is needed for the non-iterative computation by matrix inversion. The sensor geometry was as defined in Section 4.2 for two sensors. Independent drift rate fluctuations were applied in roll, pitch and yaw and μ refers to a transit, $\sqrt{\mu}$ being applied to each measurement of a pair.

TABLE 4.2: ILLUSTRATIVE ERRORS DUE TO GYRO DRIFT

(98 min. orbit with 6 mins. between star transits).
RMS Estimation errors, arc-seconds.

μ	ϕ	θ	ψ
0.95	1.40	1.89	1.38
0.90	1.06	1.08	1.02
0.80	0.81	0.66	0.74
0.70	0.53	0.54	0.65

In order to estimate performance in a 24 hour orbit the design predictions of MSDS (ref.17) have been accepted. As indicated in Table 2.1, with a sensitivity corresponding to star magnitudes 5-6, the mean crossing time would be in the range 3-9 minutes, hence the use of 6 minutes in table 4.2.

$$p(n) = \Phi(t_n - t_{n-1})p(n-1) + \Phi_i(t_n - t_{n-1})d(n-1) \quad ; \quad p(0) = 0 \quad (4.6a)$$

TABLE 4.3: ILLUSTRATIVE ERRORS DUE TO GYRO DRIFT
(24 hour orbit with 6 mins. between star transits)
RMS Estimation errors, arc-seconds.

μ	ϕ	θ	ψ
0.95	1.98	1.89	1.88
0.90	1.14	1.07	1.07
0.80	0.75	0.65	0.69
0.70	0.65	0.54	0.60

4.4

OVERALL PERFORMANCE

The results of Sections 4.2 and 4.3 can now be combined to provide representative errors in estimating satellite attitude angles by FADS using a star-crossing detector. The composite random star sensor error of 2.6 arc-seconds of equation (4.1) has been employed in conjunction with Table 4.1. Table 4.4 below lists the results in terms of $\bar{\alpha}$ as defined at (4.2), viz: RMS of roll, pitch and yaw errors.

TABLE 4.4: REPRESENTATIVE TOTAL ANGULAR ERRORS

RMS Estimation Errors ($\bar{\alpha}$ of eqn. (4.2) arc-seconds)

μ	$\bar{\alpha}_1$ (sensor)	98 min. orbit		24 hour orbit	
		$\bar{\alpha}_2$ (drift)	$\sqrt{\bar{\alpha}_1^2 + \bar{\alpha}_2^2}$	$\bar{\alpha}_2$ (drift)	$\sqrt{\bar{\alpha}_1^2 + \bar{\alpha}_2^2}$
0.95	1.42	1.57	2.12	1.92	2.39
0.90	2.36	1.05	2.58	1.09	2.60
0.80	5.90	0.74	5.95	0.70	5.94
0.70	11.96	0.58	11.97	0.60	11.98

It is evident that a long memory (μ approaching unity) is desirable to smooth the random errors of the star sensor but a short memory (μ much less than unity) permits more rapid adjustment to changes in the gyro drift rates. The results for the two kinds of orbits may be summarised as follows:

- Low-Earth orbit: optimum memory factor μ approximately equal to 0.95 (equivalent to time constant of 1.2 orbital periods) when total rms error ≈ 2.1 arc-seconds.
- 24 hour orbit: optimum memory factor μ approximately equal to 0.95 (equivalent to time constant of 1.2 orbital periods) when total rms error ≈ 2.4 arc-seconds.

Performance would improve with more frequent star transits, e.g. if a more sensitive optical system permitted detection down to magnitude 4 stars in the case of the low orbit with the higher scan rate.

5.

CONCLUDING REMARKS

The Filtered Attitude Determination System described in this paper is a promising prototype for use in future earth-pointing satellites, with one sigma accuracy the order of 2 arc-seconds for low-earth or geostationary orbits. The experimental demonstration will permit an early illustration of the new ESA interface specifications, viz: MACS or 'Modular Attitude Control Systems'. The estimating algorithm adopted is considered especially suitable for the microprocessors currently qualified for space applications, i.e. a basic 16 bit wordlength without hardware floating-point operations.

Recent tests on the Ferranti 125 gyro have enabled a more careful evaluation of drift over many hours, which is especially relevant to a star-crossing strap-down-inertial-optical system. Although a unifying simple model has not emerged, the very low-frequency behaviour is dominated by a random walk process. Nevertheless, at least for low-earth orbits, the performance of the gyro is such that it is not the serious limitation that was expected.

Future developments of FADS will probably include modules for use with star trackers, RF sensors and possibly state observers for flexible spacecraft.

6.

REFERENCES

1. 'Study of Strapdown Inertial Optical Systems (SIOS) for Future ESA Spacecraft', by RAE and BAe, Bristol, March 1981. ESA contract No. 3804/78/NL/JS(SC).
2. Noton M. 'Filtered, Attitude Determination System, the Research Programme', BAe, Bristol Report ESS/SS 1080 to RAE for Department of Industry, UK.
3. King, A.D. 'Hardware Aspects of Inertial Attitude Control Systems', ESA Conference on AOCS, Noordwijk (1977).
4. 'Gyro Noise Measurement and Analysis', by NLR and Ferranti, NLR TR79077U, ESA Contract No. 3423, 1979.
5. Dodds, S.J. 'Adaptive, High Precision, Satellite Attitude Control for Microprocessor Implementation, Automatica, vol.17, No.4, 1981, 563-573.
6. Box, G.E.P. and Jenkins, G.M. 'Time Series Analysis; Forecasting and Control', Holden Day, 1970.
7. Davis, R.T. and Gibson, W.C. 'HMOS Chips yield Environmentally Hard Computer', Defense Electronics, April 1983, 60-67.
8. Rolls, C and Allen, R, 'In-Flight Software', Issue 2, ESA Publication THS/395/CR/EH, August 1979.
9. Jude, R.J. 'System Study of an Inertial Attitude Measurement System for Earth-Pointing Satellites', ESA Conference on AOCS, Noordwijk, 1977.
10. Dyer, P and McReynolds S 'Extension of Square-Root Filtering to include Process Noise', Journal of Optimization Theory and Applications, Vol.3, No. 6, 1969, 444-458.
11. Kaminski, P.G, Bryson, A.E, and Schmidt, S.F. 'Discrete Square Root Filtering: A Survey of Current Techniques', IEEE Trans. Automatic Control, AC-16, No. 6, 1971, 727-735.
12. Albert A.E, and Gardner L.A, 'Stochastic Approximation and Nonlinear Regression', MIT Press, 1967.
13. Wilcox J.C. 'A New Algorithm for Strapped-Down Inertial Navigation', IEEE Trans. Aerospace and Electronic Systems, AES-3, No. 5, 1967, 796-802.
14. Ickes, B.P. 'A New Method for performing Digital Control System Attitude Computations using Quaternions', AIAA Journal, Vol.8, No.1, 1970, 13-17.
15. Lefferts, E.J., Markley, F.L, Schuster, M.D. 'Kalman Filtering for Spacecraft Attitude Estimation', J. Guidance, Vol.5, No.5, 1982, 417-429.
16. Palmstrom, R.E. 'FADS Gyro Performance and Testing', Communication from Ferranti to BAe, 15th April, 1983.
17. Cope P. 'Fads Performance in a 24 hour orbit'. telex from MSDS to BAe-Bristol 31 Aug 1983.

7.

ACKNOWLEDGEMENTS

The work reported in this paper was under a contract from the RAE, Farnborough (co-ordinated by A.Sanneki) and funded by the UK Department of Industry. The author acknowledges helpful collaboration with Marconi Space and Defence Systems (J.R.Crowden and P.Cope) and Ferranti Navigation Systems Department (R.Palmstrom) who supplied the recent gyro test data. Other staff engaged upon this project at British Aerospace, Space and Communications Division (Bristol) have been M.Bradford, R.Husband and C.Peat.

8

APPENDICES

8.1

ANALYSIS OF CYRO DATA

8.1.1

The averaging Process

The first results required refer to the process of averaging over an interval of time T . It will be sufficient to characterise the input as one with the auto-correlation function

$$\overline{x(t)x(t+\tau)} = a^2 \exp(-W_0 \tau) \quad (8.1)$$

and spectral density (per rad/s)

$$\left(\frac{2a^2}{\pi W_0} \right) / (1 + W^2/W_0^2) \quad (8.2)$$

The variance of the output y , for input x is

$$\overline{y^2} = \frac{1}{T^2} \int_0^T d\tau_1 \int_0^T d\tau_2 \overline{x(\tau_1)x(\tau_2)} \quad (8.3)$$

$$= \frac{a^2}{T^2} \int_0^T d\tau_1 \int_0^T \exp(-W_0 |\tau_1 - \tau_2|) d\tau_2$$

$$= \frac{a^2}{T^2} \int_0^T d\tau_1 \left\{ \int_{\tau_1}^T e^{-W_0(\tau_2 - \tau_1)} d\tau_2 + \int_0^{\tau_1} e^{-W_0(\tau_1 - \tau_2)} d\tau_2 \right\}$$

$$= \frac{2a^2}{W_0^2 T^2} \left[W_0 T - 1 + e^{-W_0 T} \right] \quad (8.4)$$

Denote by y_i the value of y in the i th interval, then by a very similar calculation we deduce

$$\overline{y_i y_{i+1}} = \frac{a^2}{W_0^2 T^2} (1 - e^{-W_0 T}) \quad (8.5)$$

and consequently from (8.4) and (8.5) the correlation between successive levels y_i is

$$\rho_1 = \frac{1}{2} (1 - e^{-W_0 T})^2 / (W_0 T - 1 + e^{-W_0 T}) \quad (8.6)$$

When

$$W_0 T \gg 1$$

(8.6) approaches the limiting form

$$\rho_1 \approx \frac{1}{W_0 T} \quad (8.7)$$

and the limit of (8.4) yields similarly

$$\overline{y^2} \approx \frac{2a^2}{W_0 T} \quad (8.8)$$

It follows that if this form of averaging is applied to a process with a correlation time greater than $1/W_0$, then the serial correlation of y_i will be negligible. Thus y_i will be a random rectangular wave, with the same mean value as the input, and the well-known auto-correlation function.

$$R(t) = \frac{2a^2}{W_0 T} \left(1 - t/T \right); \quad 0 \leq t \leq T \quad (8.9)$$

$$R(t) = 0; \quad t > T$$

The corresponding spectral density function is

$$\begin{aligned} \Phi(\omega) &= \frac{2}{\pi} \int_0^\infty R(\tau) \cos \omega \tau \cdot d\tau \\ &= \frac{2a^2}{\pi W_0} \cdot \left(\frac{\sin \frac{\omega T}{2}}{\frac{\omega T}{2}} \right)^2 \end{aligned} \quad (8.10)$$

which has bandwidth to half power

$$= 0.443/T \quad \text{Hz} \quad (8.11)$$

if T is in seconds.

8.1.2

Fitting AR and MA Models

Standard computer packages are available but, since only second-order models are in use, simpler calculations sufficed. They are specified here as a matter of convenience.

Fitting to autoregressive (AR) models is easy because it reduces to linear least-squares. Thus to fit (4.4) assuming $y(n)$ is measured perfectly, (a_1, a_2) are to be chosen to minimize

$$\sum_{i=1}^{N-2} [y(i+2) - a_1 y(i+1) - a_2 y(i)]^2$$

put

$$S_{mn} = \sum_{i=1}^{N-2} w(i+m)w(i+n) \quad (8.12)$$

and the solutions can be expressed as

$$\begin{aligned} a_1 &= (S_{12} S_{00} - S_{02} S_{01})/D \\ a_2 &= (S_{02} S_{11} - S_{12} S_{01})/D \end{aligned} \quad (8.13)$$

where

$$D = (S_{11} S_{00} - S_{01}^2) \quad (8.14)$$

If the same type of approach is applied to the moving average (MA) model (4.5), it becomes a nonlinear minimization for which a computer routine should be employed. In order to avoid this we adopted the approximate approach of fitting to the computed variance and correlations as follows. Thus from (4.5) we can show that

$$\left. \begin{aligned} \overline{y^2} &= b_1^2 + b_2^2 + b_3^2 \\ \overline{y(n)y(n+1)} &= b_2 b_1 + b_3 b_2 \\ \overline{y(n)y(n+1)} &= b_3 b_1 \end{aligned} \right\} \quad (8.15)$$

There are three nonlinear simultaneous equations to solve for the three unknowns. Substitute

$$\left. \begin{aligned} b_1 &= a \sin \theta \cos \phi \\ b_2 &= a \sin \theta \sin \phi \\ b_3 &= a \cos \theta \end{aligned} \right\} \quad (8.16)$$

where

$$a^2 = \overline{y^2}$$

leaving two equations in θ and ϕ . Substitute $t = \tan \theta$, to produce one nonlinear equation in t , etc.

8.2

ANALYTICAL DETAILS OF THE ESTIMATOR

8.2.1

Relations for Section 3.2

$$\left. \begin{aligned} h_m^T &= (-\tan \beta_m \cos \gamma_j, 1, -\tan \beta_m \sin \gamma_j) \\ k_m^T &= (\sin \gamma_j, 0, -\cos \gamma_j) \end{aligned} \right\} \quad (8.17)$$

$$b_m = \begin{bmatrix} -\tan \beta_m (c_1 c_0 - s_1 s_0) \\ 1 \\ -\tan \beta_m (c_1 s_0 + s_1 c_0) \\ -\tan \beta_m (c_1 s_0 + s_1 c_0 - 1) \\ \omega_0 t \\ -\tan \beta_m (c_1 (1 - c_0) + s_1 s_0) \end{bmatrix} \quad (8.18)$$

$$c_m = \begin{bmatrix} s_1 c_0 + c_1 s_0 \\ 0 \\ s_1 s_0 - c_1 c_0 \\ s_1 s_0 + c_1 (1 - c_0) \\ 0 \\ s_1 (1 - c_0) - c_1 s_0 \end{bmatrix} \quad (8.19)$$

$$c_1 = \cos \gamma_j, \quad s_1 = \sin \gamma_j, \quad c_0 = \cos \omega_0 t, \quad s_0 = \sin \omega_0 t \quad (8.20)$$

where $j = 1, 2$ refers to the first and second sensor.

8.2.2

Recursive form of Weighted Least-Squares Estimations

Choose x to minimize

$$\sum_{m=1}^n \mu^{n-m} [(y_m' - b_m^T x)^2 + (z_m' - c_m^T x)] \quad (8.21)$$

where $0 < \mu < 1$. The standard solution of linear regression is

$$\begin{aligned} \hat{x} &= A_n^{-1} \sum_{m=1}^n \mu^{n-m} (y_m' b_m + z_m' c_m) \\ A_n &= \sum_{m=1}^n \mu^{n-m} (b_m b_m^T + c_m c_m^T) \end{aligned} \quad (8.22)$$

In order to deduce the sequential solution, note that

$$A_n = \mu A_{n-1} + (b_n b_n^T + c_n c_n^T) \quad (8.24)$$

which permits the re-arrangement.

$$\begin{aligned} \hat{x}_n &= A_n^{-1} (y_n b_n + z_n c_n) + \mu A_n^{-1} A_{n-1} \hat{x}_{n-1} \\ &= A_n^{-1} (y_n b_n + z_n c_n) + A_n^{-1} (A_n - b_n b_n^T - c_n c_n^T) \hat{x}_{n-1} \\ \text{or} \quad \hat{x}_n &= \hat{x}_{n-1} + A_n^{-1} [b_n (y_n - b_n^T \hat{x}_{n-1}) + c_n (z_n - c_n^T \hat{x}_{n-1})] \end{aligned} \quad (8.25)$$

In order to generate the inverse also by a sequential process we make use of the matrix identity.

$$[I + P H R^{-1} H^T]^{-1} P = P - P H [H^T P H + R]^{-1} H^T P \quad (8.26)$$

Put $H = bc$ and $R = \mu I$ & rearrange as

$$[\mu P^{-1} + H H^T]^{-1} = P / \mu - \frac{1}{\mu} P H [H^T P H + \mu I]^{-1} H^T P$$

Thus if $P^{-1} = A_{n-1}$

$$B_n = A_n^{-1} = [\mu A_{n-1} + H_n H_n^T]^{-1}$$

$$\text{or } B_n = \frac{1}{\mu} B_{n-1} - \frac{1}{\mu} B_{n-1} H_n [H_n^T B_{n-1} H_n + \mu I]^{-1} H_n^T B_{n-1} \quad (8.27)$$

8.2.3

Transformation to Square Root Form

Let

$$S_{n+1} = S_n \left[\frac{1}{\sqrt{\mu}} I - e_n e_n^T / \alpha_n \right] \quad (8.28)$$

where e_n is given by (3.20) and α_n by (3.22).

Form

$$\begin{aligned} S_{n+1} S_{n+1}^T &= \frac{1}{\mu} S_n S_n^T - \frac{2}{\sqrt{\mu} \alpha_n} S_n e_n e_n^T S_n^T \\ &\quad + \alpha_n^{-2} S_n e_n e_n^T e_n e_n^T S_n^T \end{aligned}$$

but $e_n^T e_n$ is scalar, hence

$$S_{n+1} S_{n+1}^T = \frac{1}{\mu} S_n S_n^T - \mu^{-1/2} \alpha_n^{-2} (2\alpha_n - \sqrt{\mu} e_n^T e_n) S_n e_n e_n^T S_n^T \quad (8.29)$$

Substitute (3.22) into (8.29) to yield

$$S_{n+1} S_{n+1}^T = \mu^{-1} S_n S_n^T - \mu^{-1} (\mu + e_n^T e_n)^{-1} S_n e_n e_n^T S_n^T \quad (8.30)$$

Now substitute (3.18) and (3.20) into (8.30), hence

$$B_n = \mu^{-1} B_{n-1} - \mu^{-1} (b_n^T B_{n-1} b_n + \mu)^{-1} B_{n-1} b_n b_n^T B_{n-1} \quad (8.31)$$

Given this has been derived for only one measurement, the bracketed term of (8.31) being a scalar, it is exactly the same as (8.27), thus confirming that (8.28) enables us to update the square root matrix and not the full covariance matrix B_n .

Furthermore, for the case of one measurement, equation (8.25) for updating the estimate, is

$$\hat{x}_n = \hat{x}_{n-1} + B_n b_n (y_n - b_n^T \hat{x}_{n-1}) \quad (8.32)$$

but from (8.31)

$$B_n b_n = \frac{B_{n-1} b_n}{\mu} \left\{ I - \frac{b_n^T B_{n-1} b_n}{(b_n^T B_{n-1} b_n + \mu)} \right\} = \frac{B_{n-1} b_n}{(b_n^T B_{n-1} b_n + \mu)} \quad (8.33)$$

Therefore (8.32) can be written, using (8.33), as shown at equation (3.19) in section 3.3.

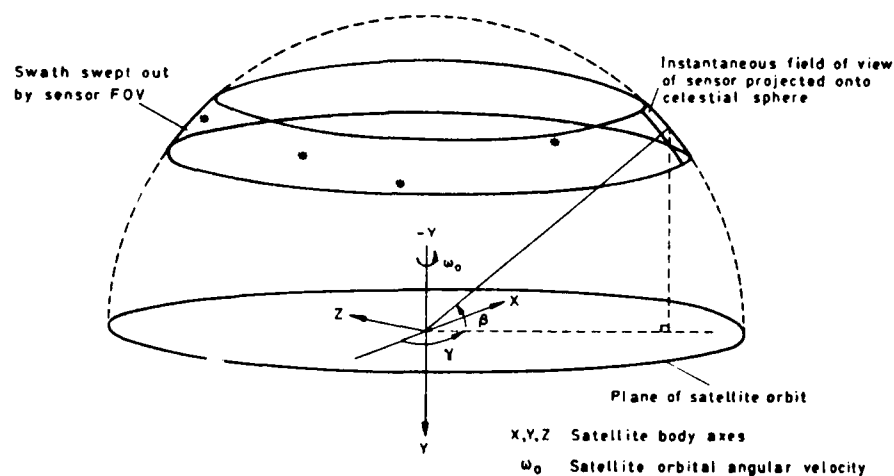


Fig. 2.1 Sensor geometry with respect to satellite body axes

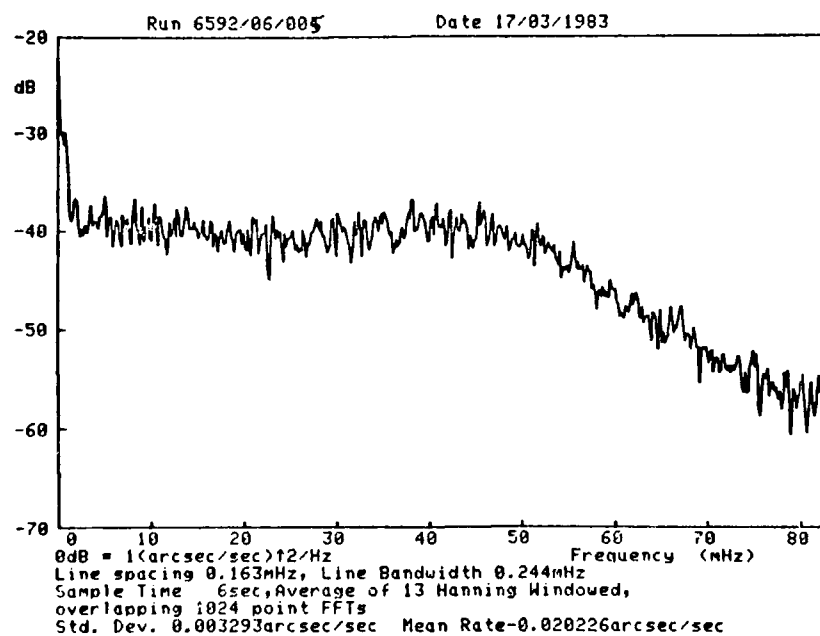


Fig. 2.2 Ferranti 125, after 6 sec averaging

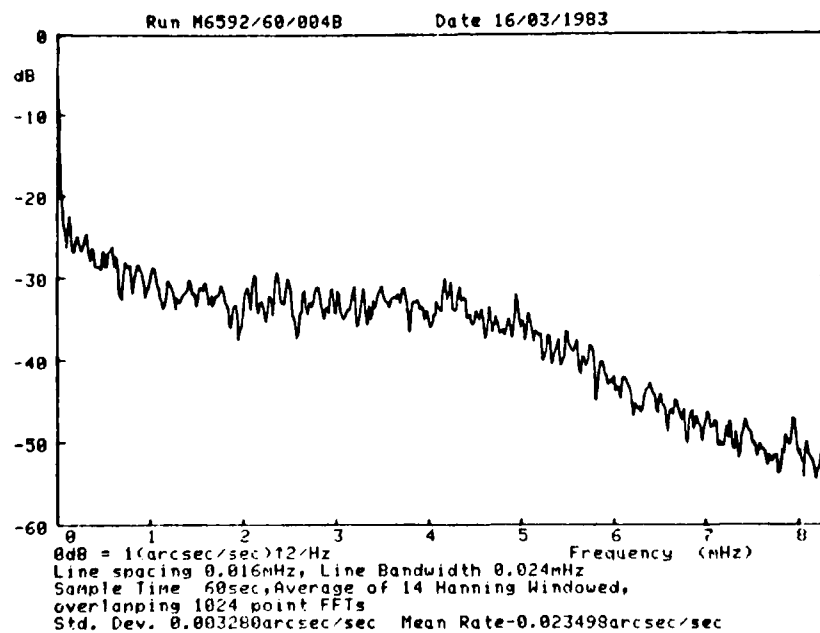


Fig. 2.3 Ferranti 125 after 60 sec averaging

FIG. 2.4 1983 DRIFT DATA ON FERRANTI 125 GYRO (RUN 50. 6 MIN INTERVALS)

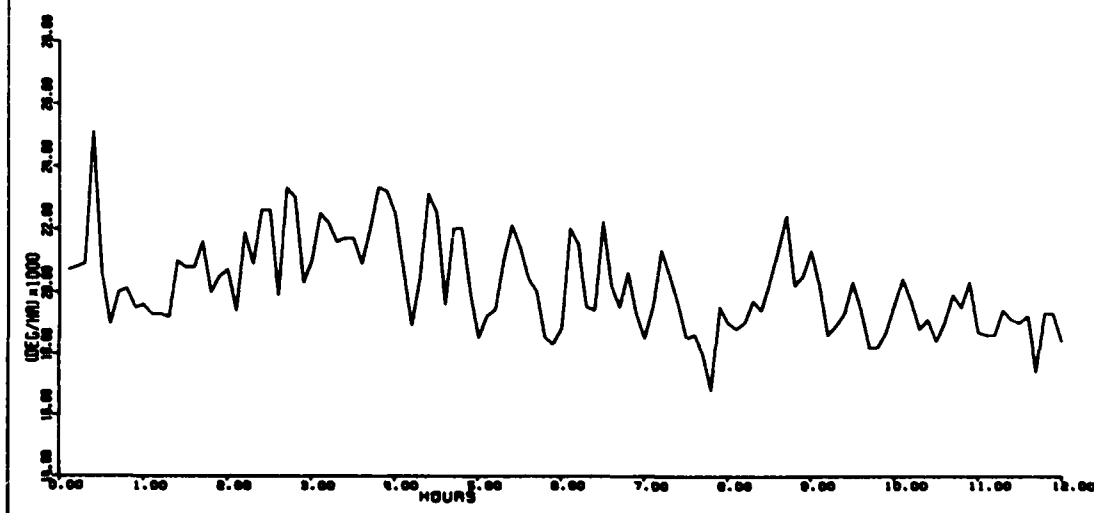


FIG. 2.5 1983 DRIFT DATA ON FERRANTI 125 GYRO (RUN 60. 6 MIN INTERVALS)

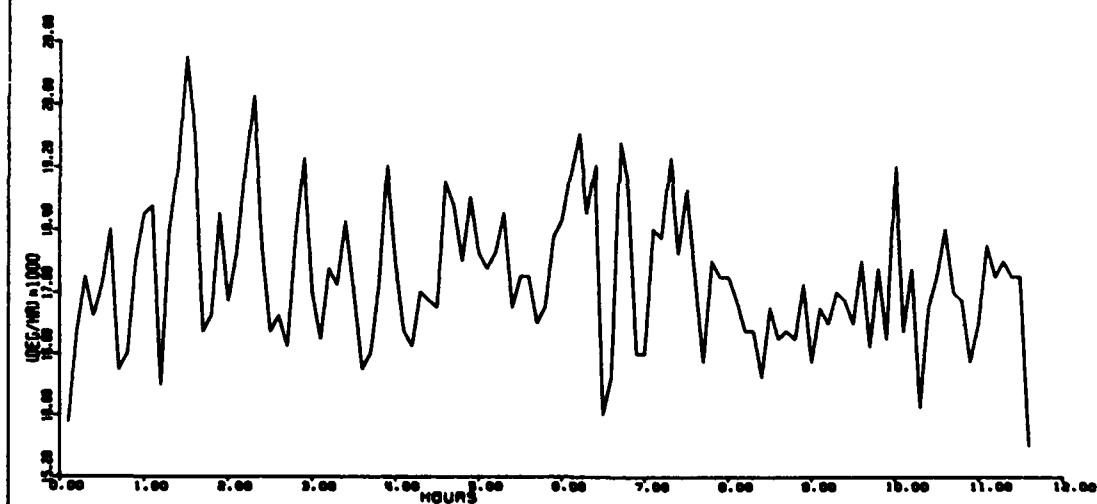
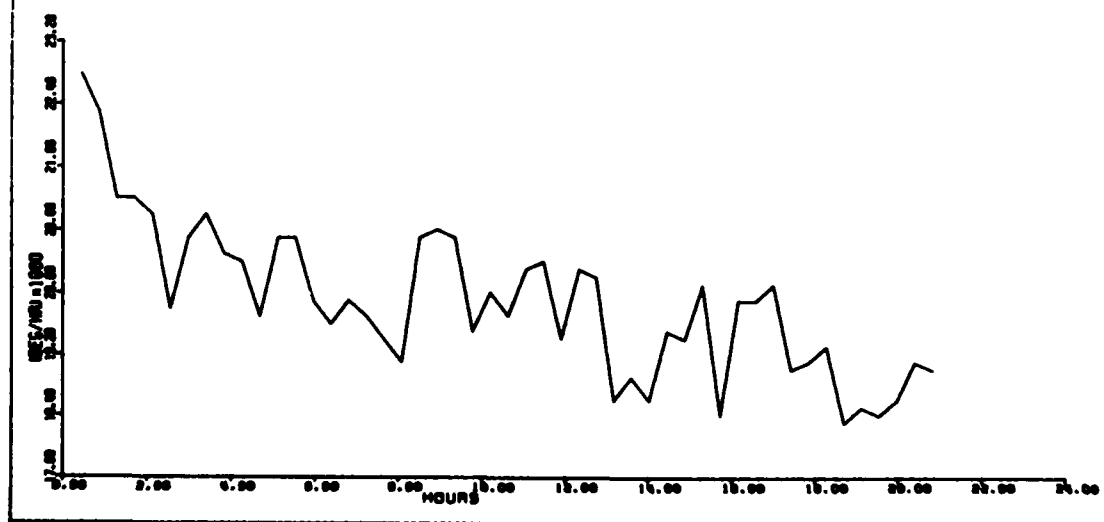
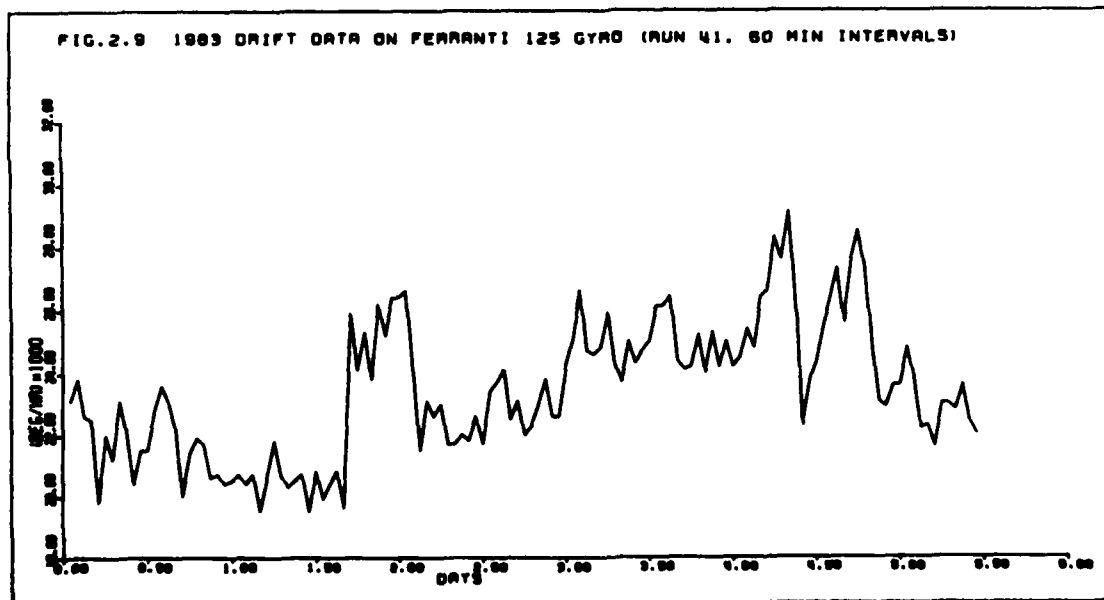
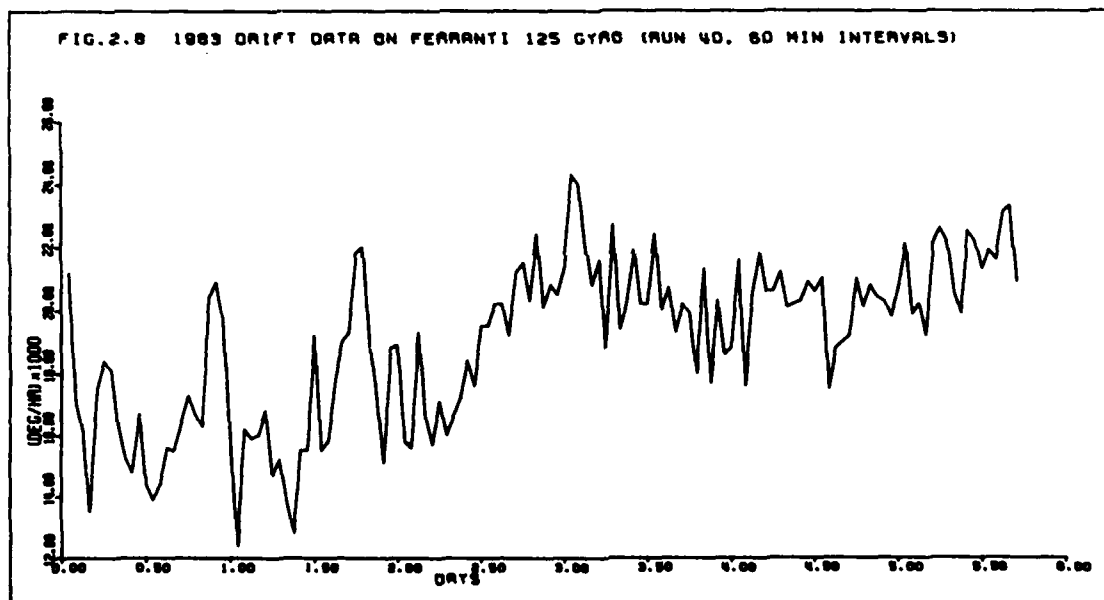
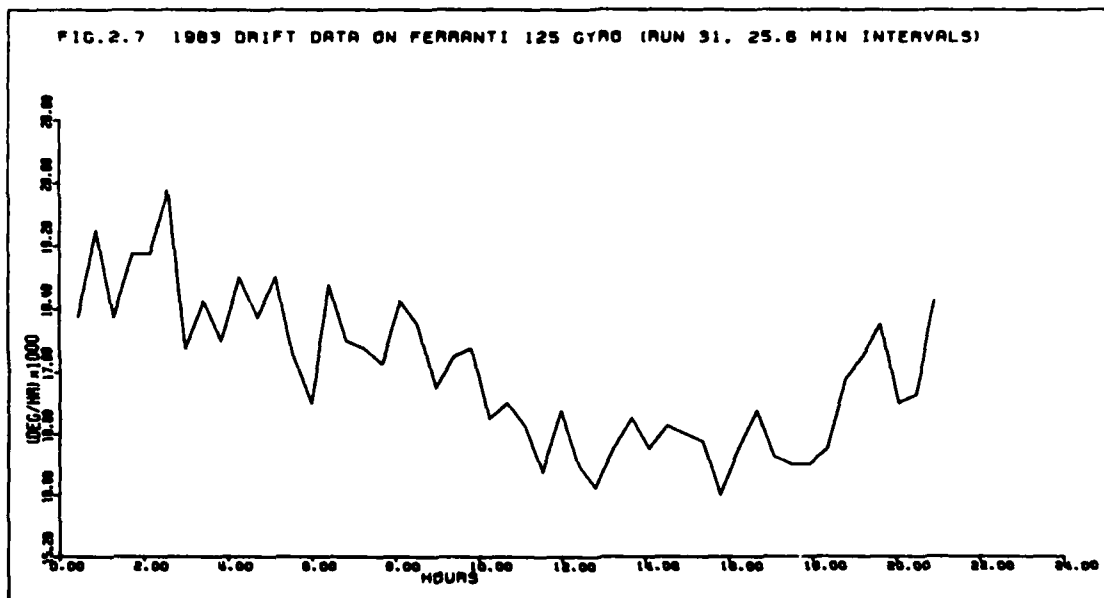


FIG. 2.6 1983 DRIFT DATA ON FERRANTI 125 GYRO (RUN 30. 25.6 MIN INTERVALS)





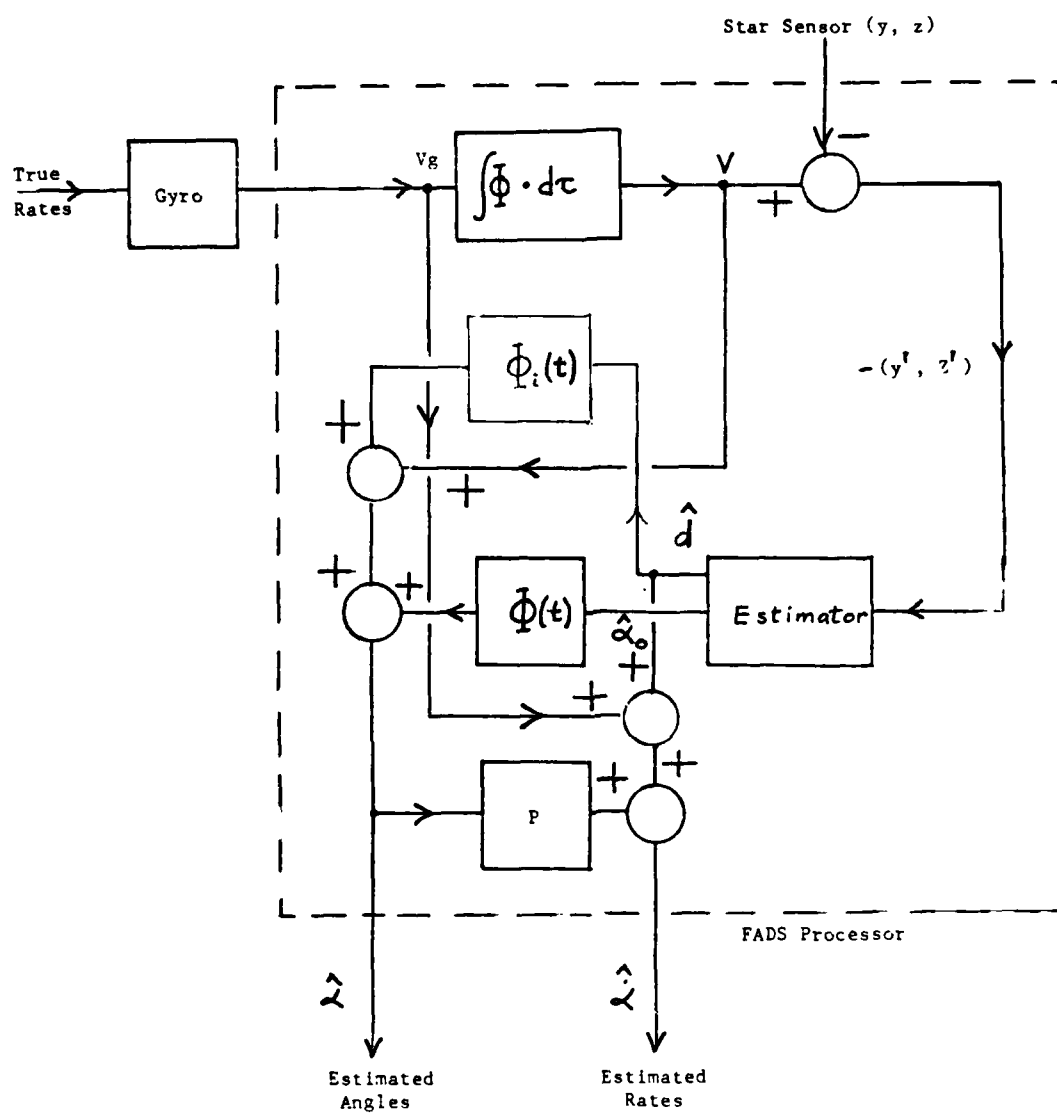


FIGURE 4.1: SCHEMATIC OF THE BASIC MODE OF FADS

INERTIAL-OPTICAL ATTITUDE DETERMINATION AND MODEL FOLLOWING CONTROL OF MANOEUVRING SPACECRAFT

by

T. Zwartbol and A.P. Terpstra

National Aerospace Laboratory NLR
Postbus 153
8300 AD EMMELOORD
The Netherlands

ABSTRACT

The paper presents an approach to on-board commanding, estimation, and control of the attitude motion of manoeuvring spacecraft, based on modern control principles.

Firstly the control algorithms are discussed. The control law is based on explicit model following. It comprises single-axis models of the spacecraft motion, which generate a target trajectory and associated target control. The spacecraft tracks the target trajectory by means of the target control which is fed forward to the spacecraft, and by means of state feedback control. A spacecraft state estimator provides for the estimates of the spacecraft state (attitude, angular velocity), and provides for the estimate of the disturbance torque which is used for disturbance torque compensation.

Secondly an algorithm for inertial-optical attitude determination and estimation of gyro parameters (drift rate bias and scale factor error) is presented.

The algorithms were validated in software simulations of an attitude control system of the type as used in the Infra Red Astronomical Satellite (IRAS). The control law and state estimator were tested in-flight, in an on-board experiment with the IRAS spacecraft.

The considered control system comprises a strapdown rate-integrating gyro and a slit-type star sensor for optical-inertial attitude sensing, a reaction wheel actuator, and a 16-bit on-board computer.

The estimation and control algorithms are described, design trade-offs are discussed, simulation results and in-flight results are presented.

1. INTRODUCTION

Several years ago the National Aerospace Laboratory NLR initiated a research programme on modern estimation and control techniques in spacecraft applications, Refs. 1, 2, 3. The present paper describes algorithms for on-board commanding, estimation and control of the attitude of manoeuvring spacecraft using an on-board computer. Their performance is demonstrated in a case study of, and in-flight experiments with the Infra Red Astronomical Satellite (IRAS, Ref. 4).

The following items are discussed:

- on-board commanding of the spacecraft attitude motion,
- a model following control technique for attitude manoeuvring,
- an estimator for state feedback and disturbance torque compensation,
- a Kalman filter for inertial-optical attitude determination and estimation of gyro parameters.

The algorithms have been tested in software simulations of a fine control system of the type as used in IRAS. The algorithms are currently being implemented on-board IRAS, and tested in flight. Some initial in-flight results are presented.

2. THE ATTITUDE CONTROL PROBLEM

The attitude control problem is depicted in figures 1, 2, 3. The satellite attitude with respect to the geocentric reference frame, figure 1, is defined by the angle ψ about the sunvector OZ, the offset angle ν about the y-axis, and the yaw angle α about the x-axis. The payload boresight is along the + x-axis. The commanded spacecraft attitude is defined by two angles (ψ_c, ν_c) where $\alpha_c = 0$ (the suffix c denotes commanded), which suffices for the payload to be able to cover the celestial sphere.

The spacecraft is to perform accurate and very stable scans and pointings about the sunvector for observation purposes. The commanded spacecraft motion consists of series of tracks of constant angular velocity $\dot{\psi}_c$ and constant offset angle ν_c ($\dot{\nu}_c = 0$). By concatenation of different tracks, defined by $(\psi_c, \dot{\psi}_c, \nu_c)$, arbitrary observation patterns can be prescribed, such as scans, rasterscans, strip-scans, pointings, etc., figure 2. The manoeuvres (slews) between the tracks, during which no observations can take place, must be performed in minimum time, figure 3.

The spacecraft is equipped with reaction wheels (RWLX, RWLY, RWLZ) along the x-, y-, and z-axis. The angles ψ and α are measured by a two-axis fine solar sensor (FSSY, FSSX). For determination and control of the attitude ψ about the sunvector a sampled, rate-integrating gyro (GYRZ) is used, whose input axis is aligned with the z-axis. A slit-type star sensor (SSE), whose line of sight is aligned along the + x-axis, provides intermittently - e.g. 3 to 12 times per orbit - additional measurements of the attitude about the sunvector. A 16-bit wordlength on-board computer (OBC) is available for data processing. The sample frequency of the control system is 2 Hz.

3. CONTROL SYSTEM DESIGN APPROACH

The equations of three-axial rotational motion of the spacecraft are coupled, nonlinear differential equations. The aim is to simplify control system design to the design of three single-axis systems, viz., the x-axis system (FSSX, RWLX) for control of $(\alpha, \dot{\alpha})$, the y-axis system (FSSY, RWLY) for control of $(\nu, \dot{\nu})$, and the z-axis system (GYRZ, RWLZ) for control of $(\psi, \dot{\psi})$. It is found from analysis of the three-axial mo-

tion, see Appendix A, that the motion $(\psi, \dot{\psi})$ can be decoupled from the other axes, where cross-coupling torques and kinematic coupling effects are small and can be considered as disturbances. However, it appears that control of the motion $(\alpha, \dot{\alpha})$ is strongly coupled with the control of the motion $(\psi, \dot{\psi})$. This is plausible, because control of the motion about the sunvector must involve simultaneous control commands to z-axis and x-axis in order to maintain $\alpha = 0$. To this end the control law for the motion $(\psi, \dot{\psi})$ about the sunvector is made to generate a control u_ψ (acceleration about the sunvector), which is distributed in the proper ratio over the z-axis and the x-axis. Likewise, a kinematic coupling $\dot{\psi}/\cos\psi$, which is a major disturbance for the motion $(\alpha, \dot{\alpha})$ during simultaneous slews, is compensated by an additional feedforward control to the x-axis. The feedforward control commands effectively decouple the equations of motion, and hence control algorithm design for $(\psi, \dot{\psi})$, and for $(\alpha, \dot{\alpha})$ and $(\psi, \dot{\psi})$ can further proceed separately.

In the next sections we limit ourselves to the discussion of the algorithms for command, estimation and control of the motion $(\psi, \dot{\psi})$ about the sunvector, on basis of the blockdiagram, figure 4.

4. SPACECRAFT MOTION COMMANDING

Figure 3 shows a generalized trajectory of the desired spacecraft motion $(\psi_c(t), \dot{\psi}_c(t))$ around the sunvector over some observation period, e.g. several hours. It consists of tracks i of constant angular velocity $\dot{\psi}_{c,i}$, beginning at initial attitude $\psi_{c,i}$, at time $t_{c,i}$.

For the purpose of commanding the spacecraft motion, a table, called Satellite Operations Plan (SOP), is loaded by ground command into the on-board computer. The SOP contains sets $\{t_{c,i}, \psi_{c,i}, \dot{\psi}_{c,i}\}$ giving the initial conditions and the desired angular velocity for each track. The commanded spacecraft state is defined by: $x_{c,k} = (\psi_{c,k}, \dot{\psi}_{c,k})^T(t_k)$, i.e. the desired spacecraft attitude and angular velocity at the control system sample points t_k .

A "setpoint generation" program in the on-board computer calculates the commanded state by linear extrapolation, according to

$$x_{c,k+1} = A x_{c,k} \quad (1)$$

where

$$A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad \Delta t = \text{sampling period of control system}$$

Algorithm (1) is properly initialized at each $t_{c,i}$, with the initial state $x_{c,i}$.

The commanded state $x_{c,k}$ constitutes the setpoint for the model following control law.

5. MODEL FOLLOWING CONTROL LAW

The following states and state errors are considered (Fig. 4):

commanded state	$x_{c,k} = (\psi_{c,k}, \dot{\psi}_{c,k})^T$
target state	$x_{T,k} = (\psi_{T,k}, \dot{\psi}_{T,k})^T$
spacecraft state	$x_k = (\psi_k, \dot{\psi}_k)^T$
target state error	$z_{T,k} = (\psi_{T,k}, \dot{\psi}_{T,k})^T = (x_{T,k} - x_{c,k})$
spacecraft state error	$z_k = (\psi_{e,k}, \dot{\psi}_{e,k})^T = (x_k - x_{c,k})$
model following error	$\epsilon_{\psi,k} = (\epsilon_{\psi 1,k}, \epsilon_{\psi 2,k})^T = (x_k - x_{T,k})$

The control law has to generate controls, such that the spacecraft state "zeroes in" on each newly commanded track in minimum time, and such that in steady-state the spacecraft state closely follows the commanded track. This is performed by an explicit model following control technique.

The control law consists - in essence - of:

- a target trajectory generator. The target trajectory generator algorithm comprises a single-axis model of the spacecraft motion about the sunvector, controlled by a time-optimal control algorithm. At each "jump" in the commanded trajectory it generates a transition trajectory or target trajectory x_T which "zeroes-in" on the commanded track x_c , figure 3. The associated target control u_T is fed forward to the spacecraft. After the transition the generated target trajectory becomes exactly equal to the commanded track and the target control vanishes until the next transition. Because the spacecraft model is noise-free, and not subject to disturbances, truly time-optimal (single-axis) transitions are generated in this way. If the spacecraft state closely tracks the target state, the spacecraft state will also "zero-in" in minimum time, and closely follow the commanded state in steady-state.
- a state feedback regulator. Due to feedforward of the target control the spacecraft tends to follow the target trajectory. However, state feedback is applied to ensure that the spacecraft state closely tracks the target state. Estimates of the attitude and angular velocity about the sunvector are produced by the spacecraft state estimator (part of the estimation algorithms, sect. 6). It also produces an estimate of the disturbance acceleration, which is used to compensate the disturbance torque.

A heuristic derivation of the discrete state equation and the components of the control law is given in the next paragraphs. Numerical values of control law parameters are summarized in table 2.

Spacecraft state equation. The control system is a sampled data system, with fixed sample rate. In Appendix A the continuous-time state equation for the motion $(\psi(t), \dot{\psi}(t))$ around the sunvector is derived, Eq. A(13), which after integration over the sample interval yields the discrete spacecraft state equation

$$x_{k+1} = A x_k + b u_{\psi,k} + b m_{\psi,k} \quad (2)$$

where

$$\begin{aligned} x_k &= (\psi, \dot{\psi})^T (t_k); \quad u_\psi = \text{the control command (acceleration) around the sunvector,} \\ &\quad \text{composed of the target control } u_T, \text{ the disturbance com-} \\ &\quad \text{pensation } u_D, \text{ and the feedback control } u_R. \\ A &= \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}; \quad m_\psi = \text{the disturbance acceleration around the sunvector.} \\ b &= (\Delta t^2/2, \Delta t)^T; \quad \Delta t = \text{the sampling interval of the control system.} \end{aligned}$$

Target trajectory generator. The spacecraft model used in the design of the target trajectory generator is the deterministic part of the spacecraft state equation (2), which is denoted by

$$x_{T,k+1} = A x_{T,k} + b u_{T,k} \quad (3)$$

where

$$x_T = (\psi_T, \dot{\psi}_T)^T \text{ is the target state.}$$

The target control of the discrete system Eq. (3) is generated by a modification of the well-known time-optimal bang-bang control law of a double integrator system (Ref. 5).

The modified, discrete minimum-time control algorithm is basically of the bang-bang type, Fig. 5.

$$\begin{aligned} u_{T,k} &= -u_{T\max} \operatorname{sign}(2\alpha_T u_{T\max} \varphi_{T,k} + \dot{\varphi}_{T,k} |\dot{\varphi}_{T,k}|), \text{ in the phase-plane} \\ u_{T,k} &= -2\alpha_T u_{T\max} \operatorname{sign}(\dot{\varphi}_{T,k}), \text{ along the switch-curve} \end{aligned} \quad (4)$$

where

$$\begin{aligned} u_{T\max} &\text{ is the maximum attainable control effort,} \\ \alpha_T &\text{ is a design parameter,} \\ \varphi_T &= \psi_T - \psi_C, \text{ is the target attitude error,} \\ \dot{\varphi}_T &= \dot{\psi}_T - \dot{\psi}_C, \text{ is the target velocity error.} \end{aligned}$$

The acceleration level during the initial phase of the transition is $u_{T\max}$, and during the terminal phase (along the switch-curve) $\alpha_T u_{T\max}$, where $\alpha_T = 0.7$ to 0.8 in order to provide a torque margin for feedback regulation and disturbance torque compensation. The other modifications account for possible constraints on the allowable spacecraft inertial angular velocity, and for the finite sampling of the system; the former by setting $u_T = 0$ (coasting) when the relative velocity limits $\dot{\omega}^+$ and $\dot{\omega}^-$ are reached, the latter by adjusting the control torque u_T upon reaching the switch-curve.

A linear dead-beat algorithm is used near the origin of the phase-plane, as the discrete bang-bang algorithm (4) will cause overshoot and repetitive switching. The deadbeat control algorithm is

$$u_{T,k} = K_{db} z_{T,k} \quad (5)$$

where K_{db} is the deadbeat gain vector. Deadbeat control is used in the region where $(0,0)$ can be reached in two sample intervals, while

$$|K_{db} z_{T,k}| \leq \alpha_T u_{T\max}. \quad (6)$$

The target control $u_{T,k}$ thus generated is fed forward to the spacecraft.

Disturbance torque compensation. Given that the target control is fed forward to the spacecraft, from Eq. (2) and Eq. (3) it appears that the disturbance torque causes the spacecraft state to deviate from the target state.

The spacecraft state estimator produces an estimate \hat{m}_ψ of the disturbance. The disturbance torque effect on the spacecraft state is counteracted by an additional control command:

$$u_{D,k} = -\hat{m}_{\psi,k} \quad (7)$$

State feedback regulator. The aforementioned control components u_T and u_D are of the open-loop feed-forward type. In order to ensure that the spacecraft indeed closely tracks the target trajectory (follows the model), the control loop is closed by a feedback regulator, producing a control command u_R . From Eqs. (2), (3), and (7) one finds that the model following error ϵ_ψ evolves according to

$$\epsilon_{\psi,k+1} = A \epsilon_{\psi,k} + b \epsilon_{m,k} + b u_{R,k}, \quad (8)$$

where

$$\epsilon_m = m_\psi - \hat{m}_\psi \text{ represents the effects of disturbance acceleration estimation errors.}$$

Feedback control is suitable to stabilize ϵ_ψ , and thence $x \rightarrow x_T$, as required.

The optimal gains of the state feedback regulator are calculated by optimization of the quadratic integral performance criterion

$$J = \int_0^\infty \epsilon_\psi^T(t) R \epsilon_\psi(t) dt. \quad (9)$$

The weighting matrix R is taken

$$R = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \quad (10)$$

where r_1, r_2 are the weights on the attitude error and the velocity error respectively. Only the ratio r_1/r_2 is important, as J can be scaled by this factor. So the relative weight r_1/r_2 can be used as a design parameter in establishing the regulator gains. Standard software packages are available for discretization of the performance index, and calculation of the optimal gains (Ref. 6).

The feedback regulator normally operates under steady-state conditions, that is, the model following error is small and hence the required control effort will not exceed the attainable maximum. However, situations can arise where larger model following errors occur, which saturate the control, i.e. the requested control exceeds the maximum control available. In general the considered control system with linear state feedback exhibits very badly damped transients when saturation of the control occurs. Therefore the feedback regulator is extended to dual mode operation (Refs. 7, 8). The final form of the feedback regulator is

$$u_{R,k} = - (k_1 \epsilon_{\psi 1,k} + k_2 \epsilon_{\psi 2,k} + k_1 \epsilon_{\psi 2,k} |\epsilon_{\psi 2,k}| / 2 \alpha_T u_{T_{\max}}); |u_{R,k}| \leq u_{T_{\max}} \quad (11)$$

where $\epsilon_{\psi 1}, \epsilon_{\psi 2}$ are the attitude and the angular velocity model following errors respectively, k_1, k_2 are the linear feedback gains, and $\alpha_T u_{T_{\max}}$ was defined previously.

For large errors Eq. (11) generates a bang-bang like, nearly time-optimal control. The function is rather steep (depending on k_1) around the "switch-curve" $\epsilon_{\psi 1} + \epsilon_{\psi 2} |\epsilon_{\psi 2}| / 2 \alpha_T u_{T_{\max}} = 0$, and the clipping of the control, $|u_{R,k}| < u_{T_{\max}}$, resembles the sign function (compare the time-optimal control law for the target model). So for large errors ($\epsilon_{\psi 1,k}, \epsilon_{\psi 2,k}$) will move towards the "switch-curve", and will move "bang-banging" around the "switch-curve" towards (0.0). Hence large errors are reduced in a nearly time-optimal manner. For small errors the quadratic velocity term is small compared to the linear velocity term, and the resulting control resembles linear state feedback.

Total control command. The total control command is given by

$$u_{\psi} = u_T + u_D + u_R \quad (12)$$

It is noted that u_{ψ} is the control acceleration for the motion around the sunvector, which must be properly distributed over x-axis and z-axis. Hence the control command to the z-axis is

$$u_{ccz} = u_{\psi} \cos \nu, \quad (13)$$

and the feedforward control command to the x-axis is

$$u_{f\psi x} = - u_{\psi} \sin \nu \quad (14)$$

6. ESTIMATION ALGORITHMS

The attitude around the sunvector is measured by means of the z-axis gyro and the star sensor. A block-diagram of the sensor data processing is given in the part of figure 4 denoted "Estimation Algorithms". The attitude around the sunvector is measured through correction and summation of the z-gyro output samples at the sample points t_k , taking into account the offset angle ν_k .

The attitude measurements from the gyro are processed further in one part of the estimation algorithms, called spacecraft state estimator, yielding estimates of the spacecraft state (attitude and angular velocity) and the disturbance acceleration.

The star sensor attitude measurements are performed intermittently, e.g. 3 - 12 times per orbit. These measurements are processed in the other part of the estimation algorithms called gyro parameter estimator, yielding estimates of the spacecraft attitude and gyro drift rate bias and scale factor error. The estimates of drift rate bias and scale factor error are used for correction of the gyro output. The attitude estimate is used for updating of the gyro attitude reference, and the spacecraft state estimator.

6.1. Measurement of the attitude around the sunvector

The gyro produces pulses representing angular increments about the spacecraft z-axis of nominally 1 arcsecond. The pulses are counted in a up-down counter, which is sampled and reset at a fixed rate (here 2 Hz).

Each gyro output sample $\Delta \omega_{g,k}$ (gyro output increment over the sample period t_{k-1}, t_k) is corrected for scale factor error and drift rate according to:

$$\Delta \omega_{gc,k} = (1 + \hat{\epsilon}_{SF}) \Delta \omega_{g,k} + \hat{\omega}_b \Delta t, \quad (15)$$

where $\Delta \omega_{gc,k}$ is the corrected gyro output sample, and $\hat{\epsilon}_{SF}$ and $\hat{\omega}_b$ are the most recent estimates of scale

factor error and drift rate bias.

The corrected "measurement" of the attitude around the sun vector then is given by:

$$\psi_{gc,k} = \hat{\psi}_1 + \sum_{i=1+1}^k \Delta\psi_{gc,i} / \cos v_i \quad (16)$$

where $\hat{\psi}_1$ is the most recent spacecraft attitude estimate established by the gyro parameter estimator.

6.2. Spacecraft state estimator

The control law uses estimates of spacecraft attitude, angular velocity, and disturbance acceleration, at the instants t_k . The estimates are obtained by the spacecraft state estimator.

Measurement equation. The spacecraft state estimator uses the integrated gyro output (Eq. 16) as its measurement. It is assumed that measurement errors - mainly quantization noise and high frequency noise - are white. Thus, the measurements can be modelled by

$$y_{s,k} = \psi_{gc,k} = \psi_k + v_{q,k} \quad (17)$$

where $v_{q,k}$ is the measurement noise, with

$$v_q(.) \sim N(0, \sigma_q^2).$$

System equation. For the design of the spacecraft state estimator the state vector from Eq. (2) is augmented by the disturbance acceleration to be estimated. The total disturbance is considered to be composed of random noise, which is modelled as a Gaussian white sequence w_{tm} , and a slowly varying acceleration, which is modelled as a random walk process m_ψ with forcing function w_m .

The resulting system equation is given by

$$x_{s,k+1} = A_s x_{s,k} + b_s u_{\psi,k} + \Gamma_s w_{s,k} \quad (18)$$

where

$$x_s = \begin{bmatrix} \psi \\ \dot{\psi} \\ m_\psi \end{bmatrix}; \quad A_s = \begin{bmatrix} 1, & \Delta t, & \frac{1}{2}\Delta t^2 \\ 0, & 1, & \Delta t \\ 0, & 0, & 1 \end{bmatrix};$$

$$b_s = \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \\ 0 \end{bmatrix}; \quad \Gamma_s = \begin{bmatrix} \frac{1}{2}\Delta t^2, & \frac{1}{2}\Delta t^2 \\ \Delta t, & \Delta t \\ 0, & 1 \end{bmatrix}; \quad w_s = \begin{bmatrix} w_{tm} \\ w_m \end{bmatrix}$$

The system noise vector w_s is a white sequence, with

$$w_s \sim N(0, Q_s), \quad Q_s = \begin{bmatrix} \sigma_{tm}^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix}$$

Estimator equations. The estimator equations follow from the Kalman filter theory

- prediction equation

$$\hat{x}_{s,k}^- = A_s \hat{x}_{s,k-1}^+ + b_s u_{\psi,k-1} \quad (19)$$

- update equation

$$\hat{x}_{s,k}^+ = \hat{x}_{s,k}^- + K_s (\psi_{gc,k} - \hat{\psi}_k^-) \quad (20)$$

where the superscripts $^-$, $-$, and $^+$ denote estimates before, respectively after the measurement update, and K_s is the estimator gain vector.

Since the system is time-invariant, the estimator gains will settle on a steady-state value. The implemented estimator uses these precalculated, steady-state gains. Standard software packages are available for calculation of estimator gains (Ref. 6).

The quantity σ_m^2 is a useful design parameter for trading-off steady-state filtering performance versus fast transient response to sudden disturbance torque changes. Disturbance torques change rapidly during manoeuvres, whereas during steady-state tracking they only vary slowly. During steady-state high filtering accuracy is required - to obtain high control accuracy - resulting in a slow transient response of the filter. During manoeuvres fast response of the filter to disturbance torque variations is required, as a slow response causes considerable transient attitude and velocity errors. In our case no acceptable compromise between steady-state accuracy and transient response speed could be found. Therefore it was decided to use two sets of gains and a gain switching mechanism. One set is used during steady-state, whereas the other set is optimized for fast transient response. Detection of sudden disturbance torque changes and switch-over to the other gains-set is performed by monitoring the measurement residual

$$r_k = y_{g,k} - \hat{\psi}_k \quad (21)$$

and comparing it with a threshold value. The switch-over criterion is given by:

$$|r_k| \geq r_{st,s} \rightarrow \text{switch-over to "fast response" gains}$$

$$|r_k| < r_{f,r} \rightarrow \text{switch-back to "steady-state" gains}$$

where $|r_k|$ is the absolute value of the current measurement residual. The threshold values $r_{st,s}$ and $r_{f,r}$ are chosen approximately equal to the (experimentally established) 3σ values of the measurement residual of the "steady state" estimator, respectively the "fast response" estimator. The numerical values of system and measurement noise and the resulting gains are summarized in Table 2.

Remarks.

- The use of the disturbance model in the estimator improves the accuracy of attitude and angular velocity estimates (cf. Ref. 13). Moreover, use of the disturbance estimate in the control law quickens the spacecraft response, and prevents steady-state attitude errors.
- The gyro measurement noise represents quantization effects and high-frequency gyro noise, which are assumed to be white. Strictly speaking there will also be other residual gyro errors, e.g. due to the limited accuracy of $\hat{\epsilon}_{SF}$ and \hat{w}_b , and residual drift. These residual errors, which may not be white, are not observable and thus cannot be filtered out. This implies that the estimates $\hat{\psi}_k$ and $\hat{\dot{\psi}}_k$ will be corrupted by residual gyro errors. The effects on $\hat{\dot{\psi}}_k$, being a low residual drift rate, are negligible. The accumulated effects in $\hat{\psi}_k$ require that $\hat{\psi}_k$ is also updated by the absolute attitude estimate $\hat{\psi}_1$, through the gyro parameter estimator (Fig. 4).
- If the spacecraft state estimator and the gyro attitude reference are updated, the attitude control system may produce a jerk which may be unacceptable from the payload point of view. In these cases the attitude update may be delayed till a more suitable time.

6.3. Gyro parameter estimator

The gyro parameter estimator is based on the system-, and measurement equations of gyroscope and star sensor. The mathematical models of these sensors and the associated system- and measurement equations are described in Appendix B.

The resulting gyro parameter estimator is of the minimum-variance Kalman-filter type. For the case considered in this study it has been assumed that 3 to 12 stars, but on the average 6 stars, can be observed per orbit. This implies that the filter will be able to estimate only the constant or nearly constant part of the gyro drift rate, i.e. the drift rate bias $w_b(t)$, and scale factor error $\epsilon_{SF}(t)$. Generally accepted models for these parameters are random walk processes to account for the long-term variations of the drift rate bias and scale factor error (Ref. 9).

The system matrix and system covariance matrix of the gyro parameter estimator are functions of the spacecraft angular velocity history and the time-interval between successive star measurements. Therefore the estimator gains in general will not settle on a steady-state value. This implies that the system matrix, the noise covariance matrix, and the filter gains have to be calculated on-board.

It is noted that $\epsilon_{SF}(t)$ and $w_b(t)$ are not observable separately if the average spacecraft angular velocity between three consecutive star measurements is constant. Thus, manoeuvres are required to obtain estimates of ϵ_{SF} and w_b separately. During periods of constant angular velocity (scan) only the drift rate bias is estimated, whereas the scale factor error is treated as a consider parameter (since the scale factor error varies slowly, one may take its uncertainty into account, but abstain from updating, (Ref. 10). Thus, scale factor error variations are estimated and corrected for via the drift rate bias estimate.

In view of the finite computer wordlength (16 bits) special attention had to be paid to numerical stability. The U-D factorization algorithm was chosen as the computationally most favourable, numerically stable algorithm (Ref. 11). Algorithmic details may be found in reference 12.

7. SOFTWARE SIMULATIONS

The performance of the estimation and control algorithms has been investigated through software simulations of an attitude control loop of a spacecraft of the IRAS type. Commanded attitude motions are pointing, scanning, and fast reorientation about the sun vector.

7.1. Truth-Models

For simulation purposes, models of the involved hardware components ("truth-models") have been developed.

The reaction wheel simulation involves torque saturation, friction, torque noise, etc.

Gyro drift and scale factor error simulation are based on discretized versions of the models given in figure 6.

Truth-model parameters are summarized in table 1. Parameter values are based on manufacturers specifications, and in-house measurements (Ref. 14 and 15).

7.2. Simulation Results

Performance of the gyro parameter estimator. The gyro parameter estimator performance is judged through the residual gyro attitude measurement error, i.e. after applying the corrections for $\hat{\epsilon}_{SF}$, \hat{w}_b , and

7. New estimates $\hat{\epsilon}_{SF}$ (if observable), \hat{w}_b , and $\hat{\psi}$ are implemented immediately after each star measurement.

Results are shown in figures 7a-b.

Figure 7a shows the performance during a calibration run in which the spacecraft is manoeuvring with both positive and negative angular velocity; both drift rate bias and scale factor error are estimated. The mean time between star measurements is 1000 sec.

Figure 7b shows the performance during a 6 orbit, constant angular velocity scan. Six stars are seen per orbit, the mean time between star measurements being 1000 sec. Because of the constant angular velocity, drift rate bias and scale factor error are not observable separately; the scale factor error effects are estimated and corrected for indirectly through the drift rate estimate.

Spacecraft state estimator performance. The spacecraft state estimator performance was studied in a number of runs, trading off steady-state filtering accuracy versus short response time to disturbance torque changes.

An example of the disturbance torque estimation response is given in figures 8a-b. It will be clear that the transient (torque) estimation errors will cause transient attitude and velocity errors. In order to obtain adequate steady-state filtering performance, as well as fast transient response, and limited attitude and velocity error transients, the estimator was implemented with two sets of gains (section 6.2).

The resulting disturbance torque estimation response, and attitude and velocity error transients are shown in figures 8b-d. It may be concluded that the estimator performance is satisfactory. The attitude error transient caused by the disturbance torque step change does not exceed 5 arcsec, thanks to the fast estimator response.

Control law performance. Control law performance can be judged from reorientation (slew) performance and from steady-state tracking performance.

- Reorientation performance is mainly determined by the target trajectory generator.
- Tracking performance is determined by the disturbance torque compensation and steady-state feedback regulation.

The "overall" tracking performance appears from figures 8c-d. The short term 1σ values of the attitude c.q. velocity error are 0.36 arcsec and 0.38 arcsec/sec respectively. Reorientation (slew) performance is shown in figures 9a-b. Reorientation manoeuvres are time-optimal within the given constraints of the target trajectory generator. This may be concluded from figure 9a, which shows the spacecraft attitude error (= actual attitude - commanded attitude) during a slew. The jump in the commanded attitude causes a jump in the attitude error, which is reduced in minimum time. It is noted that there is no transient at the end of the slew.

The slew-time equals the theoretical minimum slew-time of the target generator. It may be concluded that the spacecraft closely tracks the target trajectory and that reorientation of the actual spacecraft takes place in minimum time.

8. IN-FLIGHT RESULTS

The control law and spacecraft state estimator have been tested in flight. To this end the algorithms have been loaded from the ground in the on-board computer of IRAS, where they have temporarily taken over fine control around the sunvector from the existing PID/dual mode control law. Some results are shown in figures 10a, b, c. The commanded motion is a rasterscan, comprising two tracks, at positive c.q. negative angular velocity $|\dot{\psi}_c| = 1.925$ arcmin/sec, of 60 seconds duration each, separated by a small slew for angular velocity reversal. A transition time of 10 seconds between the tracks is taken into account, which corresponds with a jump in the commanded attitude of $\Delta\psi_c = 10 \times 1925 \times 60 = 1155$ arcsec. The optimal transition time for this jump/angular velocity reversal is $t_s = 8$ seconds.

It can be concluded from figures 10a, b that the model following control performs well during the transition. The measured attitude error ($\psi_{me} = \psi_{gc} - \psi_c$) and velocity error ($\dot{\psi}_{me} = \dot{\psi}_{gc} - \dot{\psi}_c$) are within steady-state specifications ($|\psi_{me}| \leq 5$ arcsec, $|\dot{\psi}_{me}| \leq 5$ arcsec/sec) within 9 seconds.

However, a slight overshoot (≈ 4.5 arcsec) and a transient behaviour are observed in the attitude error at the end of the slew. Although within steady-state specifications, it is a flaw compared to the simulation results (Fig. 9).

The reason for the overshoot is a mismatch of the moment of inertia of the spacecraft model - used in the target trajectory and state estimator - with respect to the actual moment of inertia of the spacecraft. It is noted that the inertia of IRAS has decreased since launch because of expenditure of the coolant of the telescope. The more oscillatory behaviour - compared to the simulation results - is due to the processing delay between the sampling instant and the issue of the new control command. The delay is 0.07 sec (fixed), the sampling interval is 0.5 sec. on board IRAS. The delay was not incorporated in the software simulations. The in-flight behaviour has been confirmed by simulations incorporating the delay and inertia mismatch. The in-flight behaviour can be further improved by reduction of the inertia mismatch, which is simply a matter of adjusting one parameter. The delay can be compensated by a simple modification of the target trajectory generator and state estimator, involving an extra prediction cycle over the processing delay. In this way the original performance can be re-achieved.

The steady-state tracking performance is very good. This is in particular apparent from the quiet control behaviour, figure 10c, and the velocity stability, figure 10b. There is no steady-state attitude error, figure 10a.

Flight results from the IRAS spacecraft, obtained with the conventional, well tuned PID/dual mode control law are shown in figures 10d,e,f. The commanded motion is a similar rasterscan. The performance improvement due to the new control algorithms appears from comparison of the 1σ values.

The results, figures 10a,b,d,e, were obtained from unfiltered gyro measurements. The gyro output is sampled every 0.5 sec., and has a quantization of 1 arcsec. Hence, the 1σ values of the measured attitude error and velocity error are larger than the 1σ values obtained from the software simulations, where the true attitude and velocity are known from the simulation.

Remarks

- Due to the good steady-state filtering performance of the spacecraft state estimator a low-noise estimate of the disturbance acceleration is obtained for cancellation of the disturbance torque.

- Manoeuvres are performed mainly in feedforward by the target trajectory generator. Therefore the bandwidth of the feedback loop can be decreased and the damping can be increased without resulting in a slow overall response.
- The application of the state estimator and state feedback enables higher damping to be achieved compared to a conventional PID control law, where the D-term causes unwanted amplification of noise.

9. SUMMARY AND CONCLUSIONS

The paper presents algorithms for on-board, sampled data, estimation and control of the attitude motion of rigid spacecraft.

The presented control law is based on model following control principles. It integrates the spacecraft reorientation problem and the steady-state tracking problem. Reorientation is performed mainly by a target trajectory generator which calculates a time-optimal transition trajectory and associated target control. Tracking of the target trajectory is performed by state feedback regulation. The feedback regulator can be optimized for steady-state tracking performance, due to feedforward of the target control and estimated disturbance acceleration.

The estimation algorithms concern the filtering of inertial-optical attitude sensor data, and spacecraft state estimation. The gyro parameter estimator uses intermittent star sensor measurements to calculate estimates of the spacecraft attitude, the gyro drift rate bias, and (if observable) the scale factor error. The estimates are used for updating the attitude reference, and online correction of the gyro output for bias drift and scale factor error. The spacecraft state estimator produces estimates of spacecraft attitude, angular velocity and disturbance torque, using the corrected gyro measurements.

Software simulations indicate satisfactory performance of the algorithms.

The control law and spacecraft state estimator have been tested in an in-flight experiment on-board the IRAS spacecraft. Reorientation (slew) performance is close to time-optimal and hence satisfactory. However, incorporation of the processing delay (between sampling and issue of the control command) in the target trajectory generator and spacecraft state estimator and better tuning of the spacecraft model inertia will further improve the response. Steady-state tracking performance is very good.

REFERENCES

1. Traas, C.R. Digital filtering methods, with applications to spacecraft attitude determination in the presence of modelling errors. Part one: Theory. NLR TR 76048 U, 1976. Part two: Extension of theory and applications. NLR TR 79039 U, 1979.
2. Woerkom, P.Th.L.M. van. Survey of modern control and observation theory for disturbed, target following aerospace systems. Part one: Basic theory. NLR TR 76041 U, 1976. Part two: Stochastic performance. NLR TR 78082 U, 1978.
3. Woerkom, P.Th.L.M. van. Discrete, bimodal control of the attitude motion of disturbed, target following spacecraft. NLR R-81-028 U, 1979.
4. Bezooyen, R.W.H. van. IRAS attitude control subsystem. Proc. of ESA conference on "Attitude and orbit control systems". Noordwijk, the Netherlands, 1977. ESA SP-128, pp. 21-27.
5. Athans, M. and Falb, P.L. Optimal control. McGraw-Hill, New York, 1966, pp. 507-514.
6. Armstrong, E.S. ORACLS. A design system for linear multivariable control. Marcel Dekker Inc., N.Y., 1981.
7. Edwards, F.G. Use of reaction wheels and nonlinear control for a satellite scanning a small celestial area. NASA-TN-D-3592, 1966.
8. Prins, J.J.M. and colleagues. A study of the application of modern control principles to IRAS. NLR R-77-025, 1977.
9. Coffman, V.D. and D.B. DeBra. Estimation of gyro parameters for experimentally developed gyro models. AIAA Paper No 75-1071, Boston, Mass., 1975.
10. Jazwinski, A.H. Stochastic processes and filtering theory. Academic Press, N.Y., 1970, ch. 8, sect. 4.
11. Bierman, G.J. Factorization methods for discrete sequential estimation. Academic Press, N.Y. 1977.
12. Dam, R.F. van den. Digital spacecraft state estimation: an algorithm based on minimum variance estimates. NLR IW-81-017 U, 1981.
13. Todman, D. and V. Wood. Study of strapdown inertial-optical attitude measurement systems. Rider no. 1 - Final report. ESA CR(P)-1051, 1978.
14. Bosgra, J.A., Reumers, J.J.M., King, A.D., and colleagues. Gyro noise measurement and analysis. NLR TR 79077 U, 1979.
15. Bosgra, J.A. and Smilde, H. Experimental and system study of reaction wheels. Part I: Measurement and statistical analysis of force and torque irregularities. NLR TR 82003 U, PART I, 1981.

ACKNOWLEDGEMENT.

The authors wish to thank their colleagues Mr. J.J.M. Prins and Mr. P.Th.L.M. van Woerkom for their contributions to the study. They also wish to thank their colleagues Mr. R.C. van Holtz, Mr. G.J. Ha-meetman and Mr. C.P.R.C. Slippens for the execution of the in-flight experiments.

Part of this study was performed under contract with the Netherlands Agency for Aerospace Programs NIVR.

APPENDIX A : DERIVATION OF THE SPACECRAFT STATE EQUATION

Reference frames. The following orthogonal right-handed reference frames are defined, figure 1:

- spacecraft centered, body-fixed reference frame $oxyz$.
- geocentric reference frame $OXYZ$.

The geocentric reference frame is treated as an inertial frame in order not to complicate the derivations unnecessarily.

Attitude angles. The attitude of the body fixed frame with respect to the geocentric frame is described by the angles ψ , ν , α . These are obtained by three successive rotations of the body fixed frame, from the initial attitude coinciding with the geocentric frame to the final attitude, in the following order: first a rotation ψ about the initial z -axis coinciding with the sunvector, secondly a rotation ν about the newly displaced y -axis, thirdly a rotation α about the final x -axis.

Commanded attitude. The commanded motion of the spacecraft is about the sunvector and about the y -axis only. Hence the commanded spacecraft attitude is defined by two angles; the angle ψ_c about the sunvector, and the offset angle ν_c , while $\alpha_c = 0$.

Kinematic relations. The relation between the angular velocities $\dot{\phi}_x$, $\dot{\phi}_y$, $\dot{\phi}_z$ about the spacecraft body-fixed axes and the rates of change of the attitude angles is:

$$\begin{vmatrix} \dot{\alpha} \\ \dot{\nu} \\ \dot{\psi} \end{vmatrix} = \begin{vmatrix} 1 & \sin\alpha \operatorname{tg}\nu & \cos\alpha \operatorname{tg}\nu \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha/\cos\nu & \cos\alpha/\cos\nu \end{vmatrix} \begin{vmatrix} \dot{\phi}_x \\ \dot{\phi}_y \\ \dot{\phi}_z \end{vmatrix} \quad (A.1)$$

The inverse relation is:

$$\begin{vmatrix} \dot{\phi}_x \\ \dot{\phi}_y \\ \dot{\phi}_z \end{vmatrix} = \begin{vmatrix} 1 & 0 & -\sin\nu \\ 0 & \cos\alpha & \sin\alpha \cos\nu \\ 0 & -\sin\alpha & \cos\alpha \cos\nu \end{vmatrix} \begin{vmatrix} \dot{\alpha} \\ \dot{\nu} \\ \dot{\psi} \end{vmatrix} \quad (A.2)$$

Differentiation of Eq. (A.1), and elimination of $\dot{\phi}_x$, $\dot{\phi}_y$, $\dot{\phi}_z$ by substitution of Eq. (A.2) in the result leads to:

$$\begin{vmatrix} \ddot{\alpha} \\ \ddot{\nu} \\ \ddot{\psi} \end{vmatrix} = \begin{vmatrix} 1 & \sin\alpha \operatorname{tg}\nu & \cos\alpha \operatorname{tg}\nu \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha/\cos\nu & \cos\alpha/\cos\nu \end{vmatrix} \begin{vmatrix} \dot{\phi}_x \\ \dot{\phi}_y \\ \dot{\phi}_z \end{vmatrix} + \begin{vmatrix} \dot{\alpha} \operatorname{tg}\nu + \dot{\psi} \dot{\nu} / \cos\nu \\ -\dot{\psi} \dot{\alpha} \cos\nu \\ \dot{\nu} \dot{\alpha} / \cos\nu + \dot{\psi} \dot{\nu} \operatorname{tg}\nu \end{vmatrix} \quad (A.3)$$

Spacecraft dynamics. The angular acceleration around the body axes $\ddot{\phi}_x$, $\ddot{\phi}_y$, $\ddot{\phi}_z$, are obtained from the spacecraft dynamics, i.e. the Euler equation.

The spacecraft is a rigid body, equipped with x -, y -, and z -axis reaction wheels, whose spin axes are aligned parallel to the body-fixed axes. It is assumed that the body-fixed axes coincide with the spacecraft principle body axes. The moments of inertia of the reaction wheels are assumed to be small compared to the moments of inertia of the spacecraft. Then the equations of motion can be linearized, yielding

$$\begin{aligned} \ddot{\phi}_x &= u_{ccx} + m'_x \\ \ddot{\phi}_y &= u_{ccy} + m'_y \\ \ddot{\phi}_z &= u_{ccz} + m'_z \end{aligned} \quad (A.4)$$

where, u_{ccx} , u_{ccy} , u_{ccz} are the commanded accelerations around the body axes, m'_x , m'_y , and m'_z are disturbance accelerations. The disturbances include: Coulomb friction, viscous friction and torque noise of reaction wheels, cross-coupling and gyroscopic torques, and external disturbances.

Spacecraft state equations. It is assumed that the deviations of α from its reference value ($\alpha_c = 0$) are small. Linearization of Eq. (A.3) around $\alpha = 0$, and substitution of Eq. (A.4) yields

$$\ddot{\alpha} = u_{ccx} + m'_x + u_{ccz} \operatorname{tg}\nu \quad (A.5)$$

$$\ddot{\nu} = u_{ccy} + m'_y \quad (A.6)$$

$$\ddot{\psi} = u_{ccz}/\cos\nu + m'_z/\cos\nu, \quad (A.7)$$

where m'_x , m'_y , m'_z are given by, see Eqs. (A.3, A.4)

$$m'_x = m''_x + m'_z \operatorname{tg}\nu + m''_x + \dot{\alpha} \operatorname{tg}\nu + \dot{\psi} \dot{\nu} / \cos\nu \quad (A.8)$$

$$m'_y = m''_y + m'_y - \dot{\psi} \dot{\alpha} \cos\nu \quad (A.9)$$

$$m'_z = m''_z + m'_z + \dot{\alpha} + \dot{\psi} \dot{\nu} \sin\nu, \quad (A.10)$$

where m''_x , m''_y , m''_z model linearization errors.

The terms m'_x , m'_y , m'_z are disturbance accelerations. During steady-state motion the disturbances are small.

Only during manoeuvres the kinematic coupling terms $(1/\cos \nu)$ and $(\sin \nu)$ and other disturbances (e.g. cross-coupling) need not be small.

Eqs. (A.5) and (A.7) deserve special attention. It is desired that the spacecraft manoeuvres around the sun-vector and the y-axis, while $\nu = 0$. From Eq. (A.5) it appears that the z-axis control acts as a disturbance which is not small, especially during manoeuvres. It can be seen that control of the motion about the sun-vector should involve coupled control commands to the z-axis and the x-axis. To this end a control law for the motion about the sunvector is to be designed (rather than about the z-axis), which calculates control commands which are distributed in the proper ratio over the z-axis and the x-axis. The feedforward to the x-axis cancels the disturbance and decouples Eqs. (A.5, A.7). Of course the x-axis must further be stabilized by a suitable control law.

Rewriting Eqs. (A.5, A.6), (A.7), taking into account the above yields the final state space equations

$$\begin{bmatrix} \ddot{x} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} m_x \quad (\text{A.11})$$

$$\begin{bmatrix} \ddot{y} \\ \ddot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{y} \\ \dot{v} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_y + \begin{bmatrix} 0 \\ 1 \end{bmatrix} m_y \quad (\text{A.12})$$

$$\begin{bmatrix} \ddot{\psi} \\ \ddot{\psi} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\psi} \\ \dot{\psi} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_{\psi} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} m_{\psi}, \quad (\text{A.13})$$

where $m_x = m_z/\cos \nu$ is the disturbance acceleration about the sunvector, and u_{ψ} is the control about the sunvector.

Here u_x, u_y, u_{ψ} are controls for which control laws are to be designed.

The total control commands around the body axes are

$$u_{ccx} = u_x - u_{\psi} \sin \nu \quad (\text{A.14})$$

$$u_{ccy} = u_y \quad (\text{A.15})$$

$$u_{ccz} = u_{\psi} \cos \nu \quad (\text{A.16})$$

where $-u_{\psi} \sin \nu$ is the feedforward control component from the control about the sunvector to the x-axis, cancelling the term $u_{ccz} \tan \nu$ and thus decoupling Eqs. (A.5) and (A.7).

Further decoupling can be achieved by additional feedforward control commands $-\dot{\psi}/\cos \nu$ to be added to u_{ccx} , and $-\dot{\psi} \sin \nu$ to be added to u_{ccz} . These controls cancel the main kinematic coupling terms, Eqs. (A.8, A.10).

If necessary, other disturbance components from Eqs. (A.8, A.9, A.10) can be cancelled by feedforward as well. In the IRAS control system both the above feedforward controls to the x-axis have been implemented, the feedforward to the z-axis has been omitted.

APPENDIX B : MEASUREMENT- AND SYSTEM EQUATION FOR THE GYRO PARAMETER ESTIMATOR

Measurement equation of a slit-type star sensor. A slit-type star sensor in essence establishes the time point t_1 at which the spacecraft attitude equals a known attitude. The star sensor measurements can therefore be modelled by the measurement equation

$$y(t_1) = v(t_1) + v_s(t_1) \quad (\text{B.1})$$

Star sensor measurement noise is modelled as a Gaussian white sequence with

$$v_s(.) \sim N(0, \sigma_{ss}^2).$$

System equation of the gyro parameter estimator. The static input-output relation of the gyro can be described by (see also Fig. 6):

$$\dot{\phi}_g(t) = (\dot{\phi}_{in}(t) - w_{dr}(t)) / (1 + s_f(t)), \quad (\text{B.2})$$

where $\dot{\phi}_{in}(t)$ is the gyro input rate, $w_{dr}(t)$ is the total gyro drift rate reduced to the input axis, $s_f(t)$ is the scale factor error, and $\dot{\phi}_g(t)$ is the output scaled to rate about the input axis.

Assuming no misalignments between the gyro input axis and the spacecraft axis gives $\dot{\phi}_{in}(t) = \dot{\phi}_z(t)$. Taking into account the offset angle ν , the angular velocity $\dot{\psi}$ around the sunvector is expressed in terms of gyro output by:

$$\dot{\psi}(t) = (\dot{\phi}_g(t)(1 + s_f(t)) + w_{dr}(t)) / \cos(\nu) \quad (\text{B.3})$$

The gyro model used for derivation of the estimator is equal to the simulation truth model, figure 6. The gyro drift rate $w_{dr}(t)$ is composed of a random walk process $w_b(t)$ to model the long-term bias variations, and an additive white noise process $w_{rn}(t)$ to model the rate noise.

The scale factor error $\epsilon_{SF}(t)$ is also modelled as a random walk process, to account for long-term variations. Mathematically:

$$\begin{aligned} w_{dr}(t) &= w_b(t) + w_{rn}(t) \\ \dot{w}_b(t) &= w_{wb}(t) \\ \dot{\epsilon}_{SF}(t) &= w_\epsilon(t) \end{aligned} \quad (B.4)$$

where the system noise vector $w(t) = \{w_{rn}(t), w_{wb}(t), w_\epsilon(t)\}^T$ is a Gaussian white noise process, with

$$w(t) \sim N(0, Q_w),$$

where Q_w is the spectral density matrix of $w(t)$.

Substitution of Eq. (B.4) into Eq. (B.3) yields the continuous-time system equation.

$$\begin{bmatrix} \dot{\psi} \\ \dot{w}_b \\ \dot{\epsilon}_{SF} \end{bmatrix} (t) = \begin{bmatrix} 0 & 1/\cos v(t) & \dot{\phi}_g(t)/\cos v(t) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \psi \\ w_b \\ \epsilon_{SF} \end{bmatrix} (t) + \begin{bmatrix} \dot{\phi}_g(t)/\cos v(t) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} w_{rn}(t)/\cos v(t) \\ w_{wb}(t) \\ w_\epsilon(t) \end{bmatrix} \quad (B.5)$$

Integration of Eq. (B.5) over the star sensor measurement interval, yields the discrete system equation for the gyro parameter estimator:

$$x_g(t_{l+1}) = A_g(\Delta T_s) x_g(t_l) + u_g(\Delta T_s) + w_g(\Delta T_s) + \lceil_g v_q(t_{l+1}) \quad (B.6)$$

where

$$x_g = \begin{bmatrix} \psi \\ w_b \\ \epsilon_{SF} \end{bmatrix}; \quad \begin{aligned} t_l &= \text{time previous measurement} \\ t_{l+1} &= \text{time present measurement} \\ \Delta T_s &= t_{l+1} - t_l = \text{star measurement interval} \end{aligned}$$

$$A_g(\Delta T_s) = \begin{bmatrix} 1, & \Delta T_s / \cos v(t), & \Delta \psi_g(\Delta T_s) \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{bmatrix}; \quad u_g = \begin{bmatrix} \Delta \psi_g(\Delta T_s) \\ 0 \\ 0 \end{bmatrix};$$

$$\Delta \psi_g(\Delta T_s) = \int_{t_l}^{t_{l+1}} \dot{\phi}_g(\tau) / \cos v(\tau) d\tau.$$

The discrete system noise is characterized by:

$$\lceil_g v_q(t_{l+1}) = \begin{bmatrix} v_q \\ 0 \\ 0 \end{bmatrix} \quad v_q \sim N(0, \sigma_q^2) \text{ is the quantization noise;}$$

and

$$w_g(t_{l+1}, t_l) \sim N(0, Q_g(t_{l+1}, t_l));$$

with

$$Q_g(t_{l+1}, t_l) = \int_{t_l}^{t_{l+1}} A_g(t_{l+1}, \tau) Q_w A_g^T(t_{l+1}, \tau) d\tau \quad (B.7)$$

It is noted that the system matrix $A_g(\cdot)$, the vector $u_g(\cdot)$ and the system noise covariance matrix $Q_g(\cdot)$ are to be calculated on-line from the gyro output $\dot{\phi}_g(\cdot)$. As the gyro output is sampled the integrals in Eqs.

(B.6) - (B.7) are approximated by summations. The gyro parameter estimator is based on system Eq. (B.6) and measurement Eq. (B.1). The measurement and system noise values used for the estimator design are summarized in table 2.

TABLE 1 Truth models parameters.

Simulation sampling frequency	2	Hz
Spacecraft, rigid, moment of inertia	646	kgm ²
Reaction wheel		
Max. attainable control torque	+ 0.2	Nm
Number of torque setpoints	+ 31	
Rate dependent friction coefficient	3.5×10^{-5}	Nm sec/rad
Coulomb friction torque	0.01	Nm
Reaction wheel torque noise		
$w_{Rtn} \sim N(0, \sigma_{Rtn}^2); \sigma_{Rtn}$	4×10^{-4}	Nm
Gyro model (Fig. 6)		
Drift rate bias (initial)	1.5	°/hr
Scale factor error (initial)	3.89×10^{-2}	(%)
Gyro output quantization	1	arcsec/pulse
Drift rate bias forcing function		
spectr. density q_{wb}	1.5×10^{-10}	(arcsec/sec ²) ² /Hz
Gyro rate noise, forcing function		
spectr. density q_{rn}	0.84×10^{-3}	(arcsec/sec) ² /Hz
Scale factor error, forcing function		
spectr. density q	5.8×10^{-10}	(%) ² /Hz
Star sensor model		
Measurement noise		
$v_s \sim N(0, \sigma_{ss}^2); \sigma_{ss}$	5	arcsec

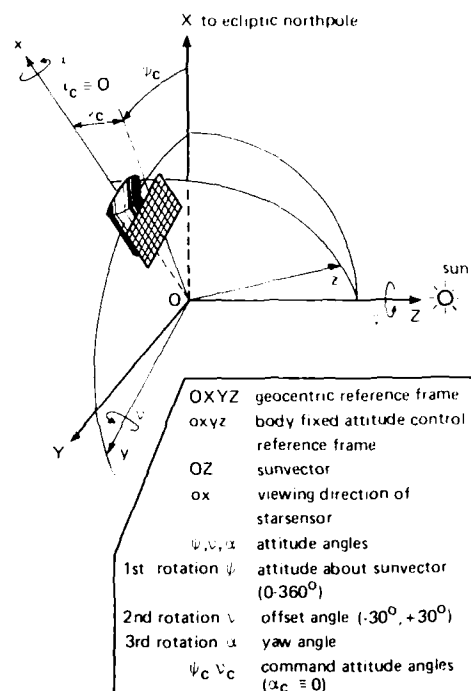


Fig. 1 Reference frames and attitude angles

TABLE 2 Control law and estimator parameters

Target trajectory generator parameters

target model moment of inertia I_T	= 646	kgm ²
max. torque	= 0.2	Nm
U_{Tmax}	= $0.2/I_T$	rad/sec ²
τ_T	= 0.8	
$K_{db} = [1/\Delta t^2, 3/2\Delta t]$	= [4, 3]	

Feedback regulator parameters

weighting matrix	$Q = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}$
LQ-optimal gains	$K_R = [0.74 \quad 2.63]$
pole-locations	0.853, -0.261

Spacecraft state estimator

gyro measurement noise $v_q \sim N(0, \sigma_q^2)$, $\sigma_q^2 = 1/12 \text{ arcsec}^2$
 system noise $w_s \sim N(0, Q_s)$

For (low) "steady state" gains:

$$Q_s = \begin{bmatrix} 3.8 \times 10^{-13} & 0 \\ 0 & 6.13 \times 10^{-14} \end{bmatrix} \begin{matrix} (\text{rad/sec}^2)^2 \\ (\text{rad/sec}^2)^2 \end{matrix} \quad K_s = \begin{matrix} 0.525 \\ 0.387 \\ 0.121 \end{matrix}$$

For (high) "fast response" gains:

$$Q_s = \begin{bmatrix} 3.8 \times 10^{-13} & 0 \\ 0 & 1.33 \times 10^{-11} \end{bmatrix} \begin{matrix} (\text{rad/sec}^2)^2 \\ (\text{rad/sec}^2)^2 \end{matrix} \quad K_s = \begin{matrix} 0.823 \\ 1.34 \\ 1.087 \end{matrix}$$

$r_{st.s} = r_{f.r} = 2 \text{ arcsec.}$

Gyro parameter estimator

Star sensor measurement noise $v_s \sim N(0, \sigma_{ss}^2)$

$$\sigma_{ss}^2 = 25 \text{ arcsec}^2$$

Gyro system noise $w(t) \sim N(0, Q_w)$

$$Q_w = \begin{bmatrix} q_{rn} & 0 & 0 \\ 0 & q_{wb} & 0 \\ 0 & 0 & q_f \end{bmatrix}; \text{ see table 1}$$

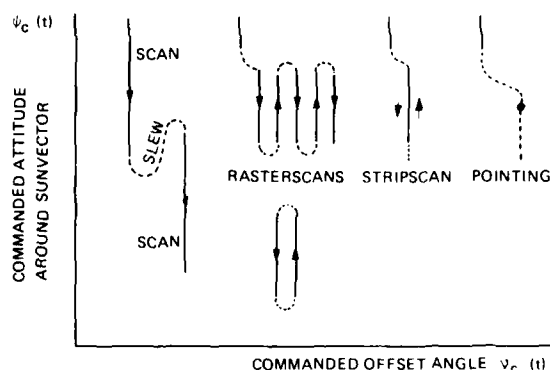


Fig. 2 Typical observation patterns on celestial sphere

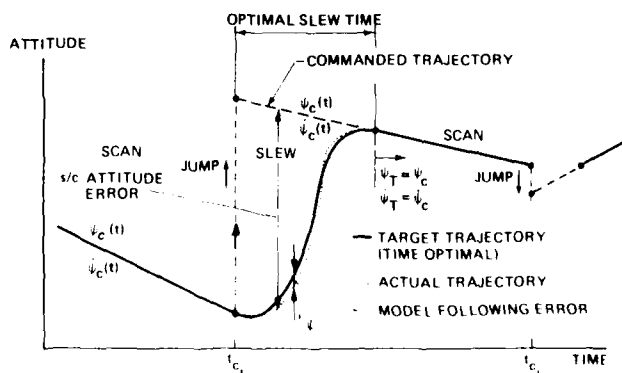


Fig. 3 Generalized commanded attitude trajectory and target - and actual trajectory, of motion about sunvector

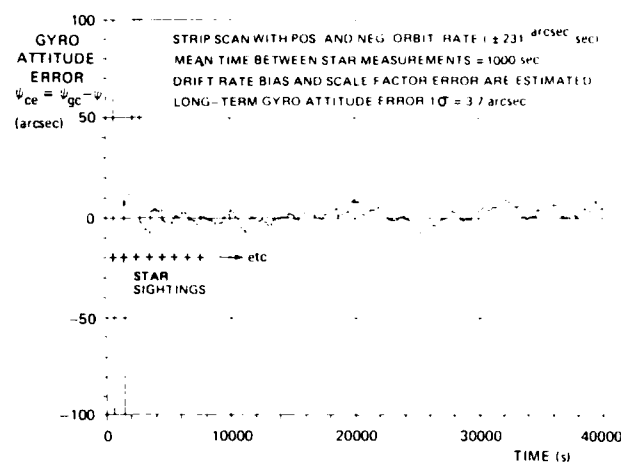


Fig. 7a Performance of gyro parameter estimator. Gyro attitude measurement error after on-line correction for estimated drift rate bias and scale factor error

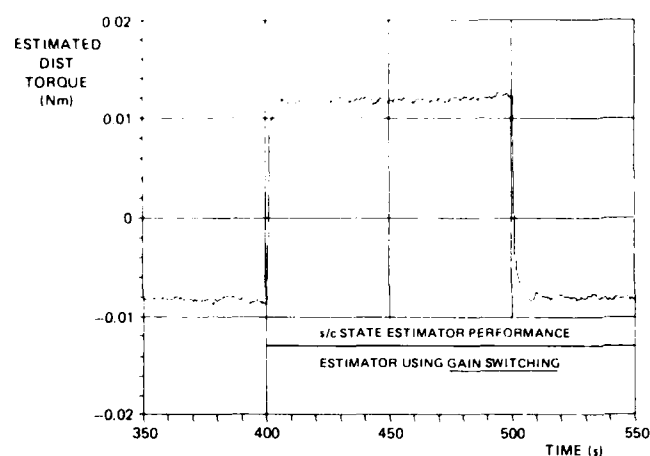


Fig. 8b Estimated disturbance torque

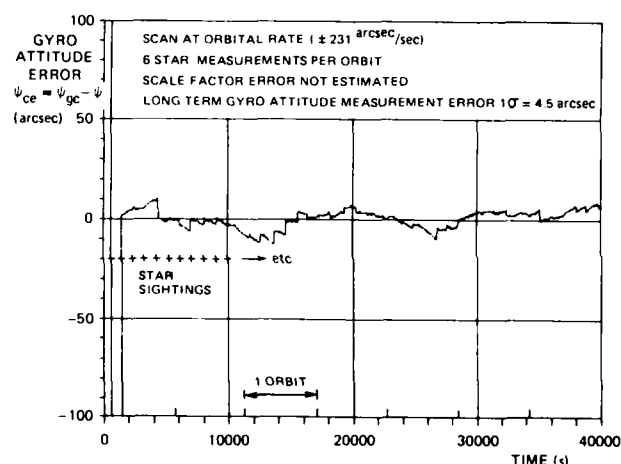


Fig. 7b Performance of gyro parameter estimator. Gyro attitude measurement error after on-line correction for estimated drift rate bias

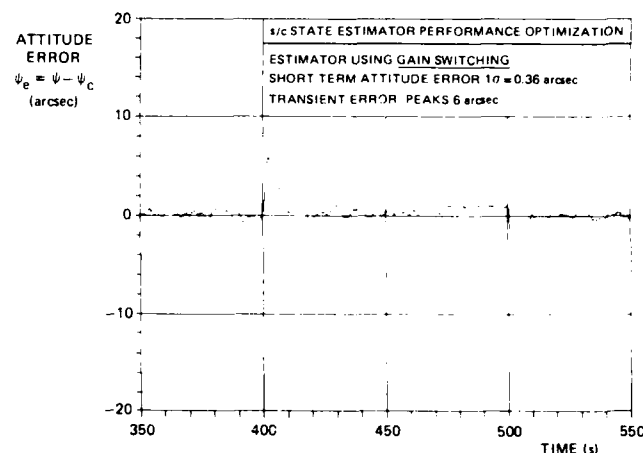


Fig. 8c Resulting control performance using estimator with gain switching: attitude error

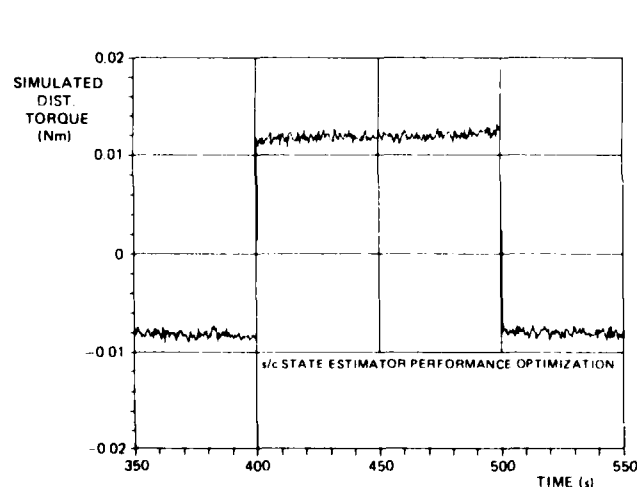


Fig. 8a Simulated disturbance torque

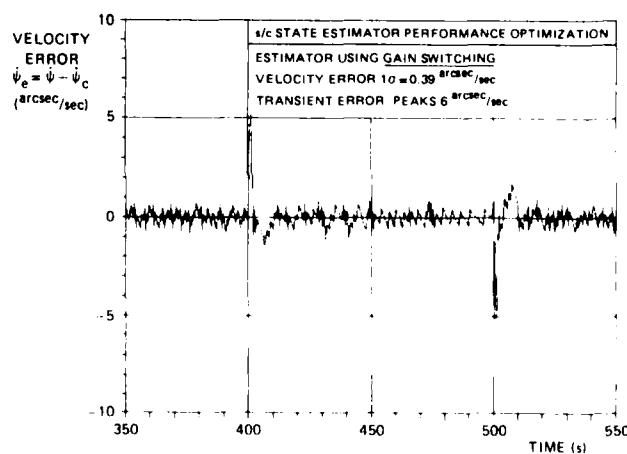


Fig. 8d Resulting control performance using estimator with gain switching: velocity error

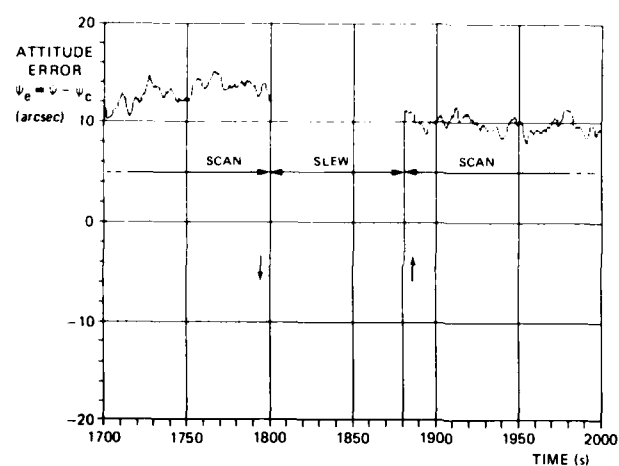


Fig. 9a Attitude error during scan-slew-scan manoeuvre

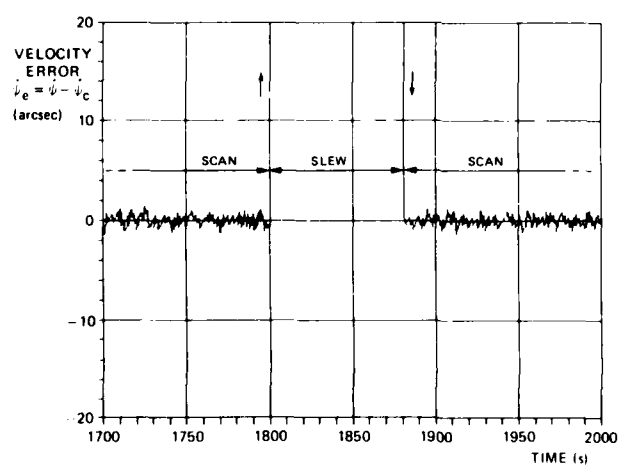


Fig. 9b Velocity error during scan-slew-scan manoeuvre

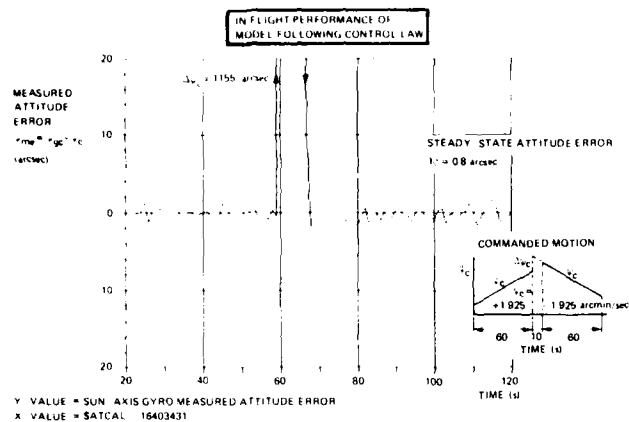


Fig. 10a In-flight attitude error

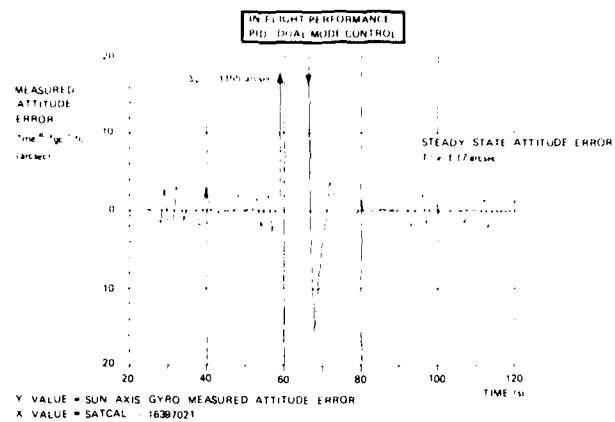


Fig. 10d In-flight attitude error

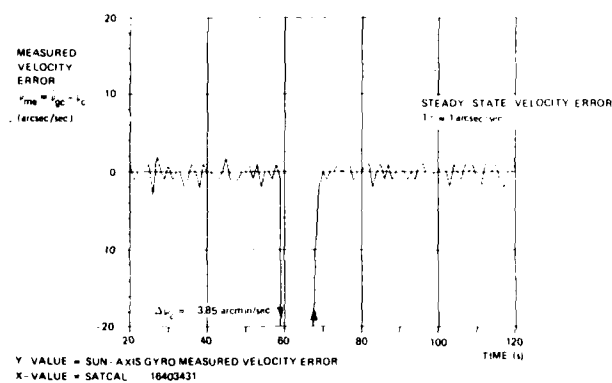


Fig. 10b In-flight velocity error

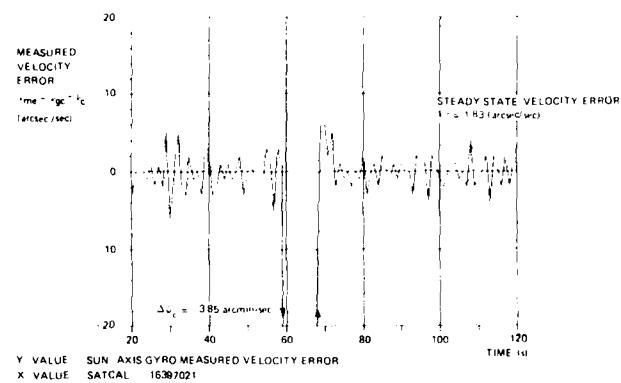


Fig. 10e In-flight velocity error

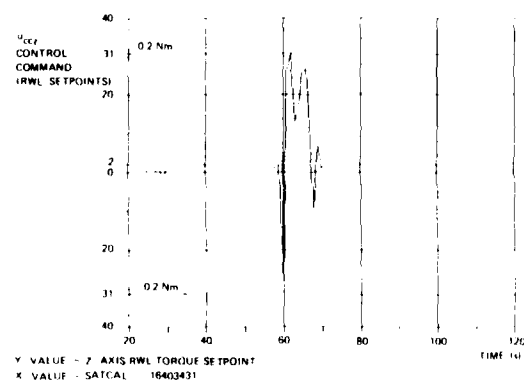


Fig. 10c In-flight control command history

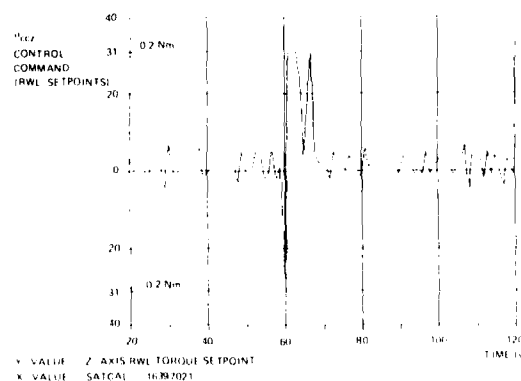


Fig. 10f In-flight control command history

DOCKING MECHANISM FOR IN-ORBIT ASSEMBLY AND SPACECRAFT SERVICING *

Dipl. Ing. Fritz Gampe
DORNIER SYSTEM GmbH
Postfach 1360
D-7990 Friedrichshafen
Fed. Rep. of Germany

SUMMARY

Future in-orbit infrastructure requires the usage of a Docking Mechanism Subsystem (DMS) for in-orbit assembly and spacecraft servicing. As a first entry for European applications unmanned, autonomous docking missions are foreseen as well in low earth orbit (processing platforms) as in geostationary orbit (communication satellite clustering). The Rendezvous and Docking (RVD) operation is a key technology to realize such type of mission.

The DMS is imbedded in a system performing the "last meters" problem, where specially the Attitude and Orbit Control Subsystem (AOCS) closely cooperates with the DMS. The DMS itself is composed of the Docking Mechanism itself and the Docking Mechanism Electronics. Depending on the initial separation conditions two docking concepts are considered: the DMS- and the AOCS - controlled closure.

A first docking demonstration is envisaged during the EURECA-2 mission in 1989/90.

1. BACKGROUND

The review of potential European missions through the next two decades shows an evolution of space use to achieve the following goals:

- extension of commercial use of space
- extension of scientific research by space means

The operations necessary to achieve these goals are summarized in Table 1-1.

The critical technology to make all these methods workable is best known under the wording. These methods involve Rendezvous and Docking (RDV), and Servicing.

According to that, RDV and servicing technology has to be developed in Europe.

(Table 1-2). This paper deals primarily with the Docking Mechanism, certain aspects of the RDV sensors, and the reliability of the total system.

2. THE RVD PHASES

The RVD process itself is composed of a sequence of RVD phases aiming at guiding an Active Chaser Satellite to a Passive Target Satellite, ending in the final mating of these two spacecrafts.

The RVD phase are:

- Inter Spacecraft Acquisition, which means the identification of the target by the chaser for navigation of the subsequent phase.
- Homing, which means that both spacecrafts are in similar orbits but the chaser Satellite being on some collision course for the other.
- Final Approach, which means that the spacecrafts are in close vicinity and the Chaser Satellite is braking-off its collision course.
- Proximity Operations, which means that the chaser is within the safety range and has to respect the safety conditions for docking, standby, flyby, retreat.
- Docking Phase, which means the docking process starting with the first physical contact, then grappling, closure and finally latching.
- Mating of the utility connectors for power and data exchange and gas/fluid supply.
- Undocking, which means, that utility mating and docking phase are reversed and the spacecrafts are separated from each other.

The Docking Mechanism Subsystem performs all operations which start with the Proximity Operation, cover the complete docking phase, and the mating of the utility connectors. To do this, the DMS is responsible to perform the following tasks of joining two spacecrafts.

- Check-out: verify DMS operational readiness for docking, and observe safety restrictions
- Grappling: make first physical contact and make firm mechanical connection between the two vehicles
- Closure: reduce kinetic energy of relative motion, align, and close vehicles together
- Latching: make final alignment and provide rigid mechanical connection
- Connection: provide electrical, gas, and fluid supply lines for inter satellite servicing.

ADP 003 390

- ASSEMBLY OF SPACECRAFTS, PLATFORMS AND ANTENNAS OUT OF MULTIPLE MODULES EXCEEDING A SINGLE LAUNCHER CAPABILITY
- OPTIMISING OF SPACE SYSTEMS AND MISSIONS DUE TO BUILT-IN FLEXIBILITY AND MODULARITY
- EXCHANGE, TRANSPORT AND RETRIEVAL OF EQUIPMENT, SAMPLES, MATERIAL, PAYLOADS, INSTRUMENTS TO AND FROM PERMANENT PLATFORMS
- REPLENISHMENT AND TRANSFER OF CONSUMABLES
- MAINTENANCE AND REPAIR OF DISABLED SPACECRAFTS
- IN-ORBIT MANIPULATION TO ADOPT A SPACECRAFT TO CHANGING MISSION REQUIREMENTS
- RE-ORBITING OF SPACECRAFT TO PROPER ORBIT

TABLE 1.-1 OPERATIONS FOR FUTURE SPACE MISSIONS

RVD TECHNOLOGY

- standardization and automation of final approach
- target/pattern recognition and image processing
- laser and/or radar systems for range, range rate and angular measurement
- multi-processing for system autonomy
- use of relay satellite(s)
- ground network involvement
- docking mechanisms (modular, mission adaptable)
- close-up sensors
- ultra reliable processing facility

SERVICING TECHNOLOGY

- Pattern recognition and image processing
- Sensing techniques and feed-back of sensor signals
- manipulation algorithms
- Remote manipulation and automation of in-orbit development
- Adaptive intelligent systems definition and development
- Manipulator mechanization
- In-orbit exchange and system re-verification
- In-orbit re-supply/fuel transfer
- Transfer technology for utilities and resources exchange

TABLE 1.-2 TECHNOLOGIES FOR DOCKING AND SERVICING IN ORBIT

3. GENERAL CONCEPT CONSIDERATIONS

3.1 Docking Concept Parameters

The concept parameters determine the design of the Docking Mechanism Subsystem. Especially "man rating", "type of docking", "control philosophy", and "mode of stabilization" are of major importance. For the European missions to which the DMS will be designed, the following parameters apply:

- unmanned vehicle
- non-impact docking
- reversible docking
- modular design
- active units concentrated on one vehicle
- autonomously controlled
- three axis stabilized
- central design
- design with possibility to add a manipulator

3.2 Docking Strategies

If during docking, the Attitude and Orbit Control Subsystem (AOCS) of a least one satellite is turned off, the docking strategy is based on a stand-off concept, where both spacecrafts are separated from each other by a noticeable distance at first physical contact. In this case we have a

DMS controlled closure

An other docking strategy, incorporating the AOCS, positions the two vehicles in very close proximity in the range of a few centimeters at first physical contact. Because of the importance role of the AOCS, we call this strategy

AOCS controlled closure

The DMS controlled closure design (see fig. 3.2-1) uses:

- a boom as stand-off device.
The boom is stowed during non-operational conditions and is extendable and retractable
- a grapple mechanism for first mechanical linkage, including the capability of undocking
- and structural latches for rigid mechanical connection, with the capability to unlatch.

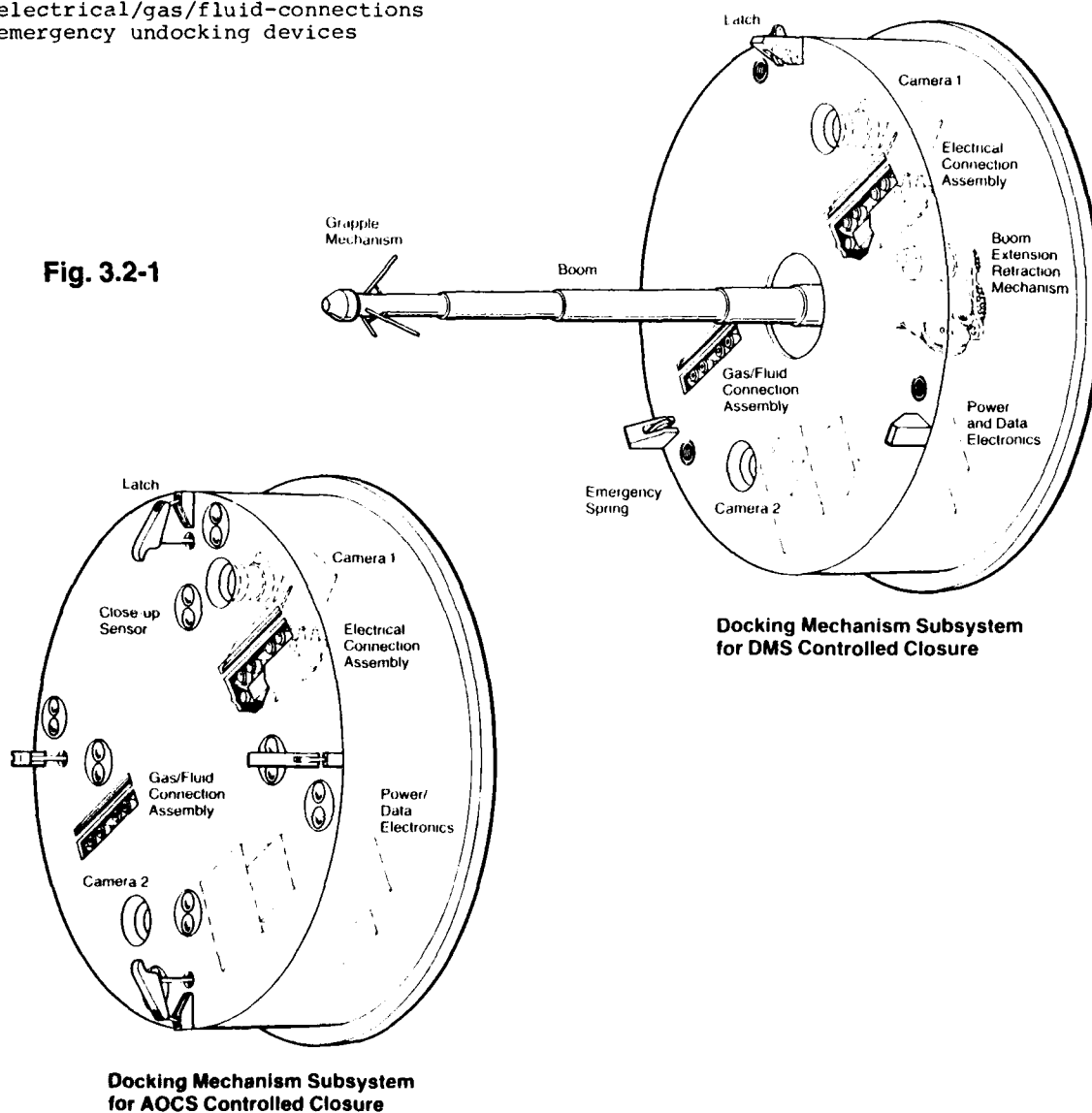
The AOCS controlled closure design (see fig. 3.2-1) uses:

- the structural latches as well for the first mechanical linkage as for the final rigid mechanical connection
- a set of near range sensors for the proximity operations indicated here as a set of video cameras for a stereo image and a set of close-up sensors indicating relative attitude measuring devices capable of measuring down to a few centimeters and for exact positioning of the two satellite surfaces.

Common features in both designs are:

- electrical/gas/fluid-connections
- emergency undocking devices

Fig. 3.2-1



4. REQUIREMENTS

After the general concept considerations the major requirements for the DMS design are presented here.

The Operational Requirements are shown in Table 4-1.

The Operational Principles state, that unmanned vehicles, are concerned and, at least for the last meter problem, do not involve the man in the control loop.

The Operational Modes are indicating the duration of the docking period

- permanent docking stands for the intention to dock once and leave the two vehicles together
- episodic docking stands for the intention of a multiple docking capability.

Nominal Operations describe the actions to be taken during a nominal performing docking process.

Contingency Operations describe how to cope in general with unforeseen events which might endanger either satellite.

Operational Modes:	<ul style="list-style-type: none"> - permanent docking - episodic docking
Operational Principles:	<p>The DMS will be used on unmanned space vehicles. Man involvement shall be limited to:</p> <ul style="list-style-type: none"> - Supervision of DMS operation - Interpretation of housekeeping data - Specially assigned stop/go commands - Contingency and Emergency control
Nominal Operation:	<ul style="list-style-type: none"> - The DMS shall be checked for docking readiness prior to the initiation of docking. - Nominal operations shall be based on a predetermined operational sequence. - Prior to each sequence and after each sequence go-ahead checks shall assess the status of the docking process and the DMS itself. - The DMS shall provide automatic correction and switch-over commands and/or control of those functions from the ground, according to the mission requirements.
Contingency Operations:	<ul style="list-style-type: none"> - Emergency operations shall be initiated when any system of either spacecraft is endangered and safety is no longer guaranteed. - Emergency operation shall be initiated by the DMS and/or from the ground, according to the mission requirements.

TABLE 4.-1 OPERATIONAL REQUIREMENTS

The Safety Requirements are shown in Table 4-2.

The General Safety Aspects give the standards as well for the operations, the design, and the command structure.

The Safety during Docking Operations cover the real docking process and has an abort capability at any point.

The Safety during Contingency Operations state, that this shall be done by fail-safe means.

The Safety during Emergency Operations leaves the Passive Satellite damaged so that a subsequent docking attempt is feasible.

GENERAL SAFETY ASPECTS

The DMS shall be designed and fabricated in such a manner that during operations and ground handling all hazards to personnel and other equipment are avoided, or minimized and controlled.

The DMS shall have no credible single point failure which results in an unsafe condition for either vehicle.

Safety shall be assured for all operational phases of the DMS.

The DMS shall make available provisions for docking abortion at any time and satellite release without damage to either satellite.

The DMS shall be protected against false commands.

The DMS shall be designed to a fail-safe, fail-safe standard.

The first point of contact shall be grounded.

SAFETY DURING DOCKING OPERATIONS

During the total docking process no damage shall occur to either spacecraft.

The docking operation shall be man supervised.

There will only be a man-initiated docking attempt if a redundant system within the DMS has to be used.

The docking process shall be abortable at any point within the process.

The operational performance of either vehicle shall not be violated, following docking.

SAFETY DURING CONTINGENCY OPERATIONS

The DMS shall provide fail-safe means for Contingency Operations.

SAFETY DURING EMERGENCY OPERATIONS

In the event of an emergency undocking there shall be no necessity for a redocking attempt with the same Active Satellite.

TABLE 4.-2 SAFETY REQUIREMENTS

The Initial Separation Requirements, given in Table 4-3 are the basic design data for the Docking Mechanism. Of course, the initial separation requirements reflect the two Docking Modes. These values apply to the relative movements of two reference points at each spacecraft. Limit cycle motions of the reference points have to be considered additionally.

Relative Movement	DMS Controlled Closure	AOCS Controlled Closure
Misalignment (m)	$d_{xs} \leq 0.1$ $d_{ys} \leq 0.1$ $d_{zs} \leq 1.0$	$d_{xs} \leq 0.02$ $d_{ys} \leq 0.02$ $d_{zs} \leq 0.1$
Approach velocity (cm/sec)	$v_{xs} \leq 0.5$ $v_{ys} \leq 0.5$ $v_{zs} \leq 1.5$	$v_{xs} \leq 0.1$ $v_{ys} \leq 0.1$ $v_{zs} \leq 0.3$
Angular misalignment (deg)	<p>pitch $\phi_{xs} \leq 5$</p> <p>yaw $\phi_{ys} \leq 5$</p> <p>roll $\phi_{zs} \leq 2$</p>	<p>pitch $\phi_{xs} \leq 1.0$</p> <p>yaw $\phi_{ys} \leq 1.0$</p> <p>roll $\phi_{zs} \leq 0.5$</p>
Rotational overspeed (deg/sec)	$\omega_{xs} \leq 0.1$ $\omega_{ys} \leq 0.1$ $\omega_{zs} \leq 0.05$	$\omega_{xs} \leq 0.05$ $\omega_{ys} \leq 0.05$ $\omega_{zs} \leq 0.05$

TABLE 4-3 INITIAL SEPARATION REQUIREMENTS

The Physical Properties Requirements, given in Table 4-4 describe the two spacecraft involved in the docking process.

The Initial Separation and the Physical Properties Requirements cover a very large range with which especially the mechanical design has to cope. It is obvious that this cannot be taken by one mechanism design only unless the smaller satellites would have to pay a mass penalty. Therefore, the DMS concept has to be a modular one, adaptable to the specific requirements of each individual mission.

Physical Properties of the Active Satellite		Physical Properties of the Passive Satellite	
mass (Kg)	$200 \leq m_A \leq 4000$	mass (Kg)	$1200 \leq m_p \leq 15000$
moment of inertia (Kg.m ²)	$100 \leq J_{XA} \leq 11000$	moment of inertia (Kg.m ²)	$5000 \leq J_{xp} \leq 250000$
	$100 \leq J_{YA} \leq 11000$		$9000 \leq J_{yp} \leq 520000$
	$100 \leq J_{ZA} \leq 8000$		$12000 \leq J_{zp} \leq 600000$
centre of gravity (m) (relative to DMS)	$0.05 \leq X_{CGA} \leq 0.25$	centre of gravity (m) (relative to docking port)	$0.2 \leq X_{CGP} \leq 10.0$
	$0.05 \leq Y_{CBA} \leq 0.25$		$0.2 \leq Y_{CGP} \leq 4.8$
	$0.75 \leq Z_{GA} \leq 2.00$		$0.2 \leq Z_{CGP} \leq 2.5$
eigenfrequencies (Hz)	$f_1 \geq 10 \text{ Hz}$	eigenfrequencies (Hz) ($f_1 - f_6$: rigid)	$0.9 \leq f_{7p} \leq 2.0$
	$f_2 \geq 35 \text{ Hz}$		$1.20 \leq f_{8p} \leq 2.0$
			$1.25 \leq f_{9p} \leq 2.0$

TABLE 4-4 PHYSICAL PROPERTIES REQUIREMENTS

The Connection Requirements are listed in Table 4-5 and cover the requirements for electrical (power and data), gas, and fluid utility connections.

ELECTRICAL CONNECTOR REQUIREMENTS

Power Transfer of 2 KW at a voltage of 50 V

Power Transfer of 1 KW at a voltage of 28 V

Low Rate Signal Transfer -- a total of 100 lines (AWG 24)

High Rate Data Transfer -- 100 Mb/sec

No connector shall have more than TBD (50) pins.

The High Frequency power density per connector shall be limited to TBD W average for the connector.

The mixing of signal and power lines shall be avoided where feasible.

Redundant circuits shall not use the same connector.

Emergency power shall be provided through separate connectors.

GAS CONNECTOR REQUIREMENTS

A high pressure gas connector shall be provided to supply 80 kg of gas at a rate of 100 bar/hr at an initial pressure of 280 bar, leakage TBD.

A low pressure gas connector shall be provided to supply 100 dm³ of gas at a flow rate of 100 dm³/hr at an initial pressure of 1 bar, leakage TBD.

FLUID CONNECTOR REQUIREMENTS

A fluid connector shall be provided to supply 100 kg of freon at a flow rate of 360 kg/hr against a pressure drop of 6 bar, leakage 0.1 cm³/hr.

A fluid connector shall be provided to supply 500 kg of bipropellant (TBS) at a flow rate of 20 kg/hr against a pressure of TBD bar; leakage TBD.

TABLE 4.-5 CONNECTION REQUIREMENTS

5. DOCKING MECHANISM SUBSYSTEM DESIGN

5.1 Docking Mechanism for DMS Controlled Closure

The Initial Separation Requirements dictate on the Active Satellite

- a stand-off device for bridging the stand-off distance
- a grapple mechanism for initial physical contact and firm fixture and on the Passive Satellite
- a (passive) guidance for the grapple mechanism.

This overall Docking Mechanism concept leads to the following general description for the DMS controlled closure (see Table 5.1-1).

ZERO-IMPACT, DROGUE/PROBE SYSTEM		
GENERAL DESCRIPTION		
FUNCTIONS	ACTIVE SATELLITE ELEMENTS	PASSIVE SATELLITE ELEMENTS
Grappling	<ul style="list-style-type: none"> - Grapple mechanism on top of an extendable/retractable boom for initial contact at stand-off distance and initial firm contact with alignment capability - axial and lateral dampers to handle the translational and rotational excess energy 	Passive guidance for grapple mechanism (no active parts on passive satellite)
Damping		
Closure	<ul style="list-style-type: none"> - Retraction of boom - Latches with final adjustment capability to bear the structural loads - Add-On features, independent of boom and latches - Opening of latches, extension of boom and release of grapple mechanism - Pyrotechnics Devices - Springs to guarantee minimum separation velocity 	OPTIONAL ELEMENTS
Latching		<ul style="list-style-type: none"> - Actuation capability of boom to minimize initial misalignment - active damping by use of specific elements and control software
Connection		
Undocking		
Emergency Provisions		

TABLE 5.1-1: GENERAL DESCRIPTION OF DMS CONTROLLED CLOSURE CONCEPT

The general outline of that system is given in Fig. 5.1-2 and the detailed is given in Fig. 5.1-3. This concept, as shown in Fig. 5.1-3 tries to incorporate at a very early stage of the conceptual design the safety and redundancy features.

The conceptual design features:

- a telescopic boom as stand-off device
- a grapple latch with three legs
- a redundant grapple latch drive mechanism to open and close the legs
- a passive drogue on the passive satellite with grooves for roll alignment.

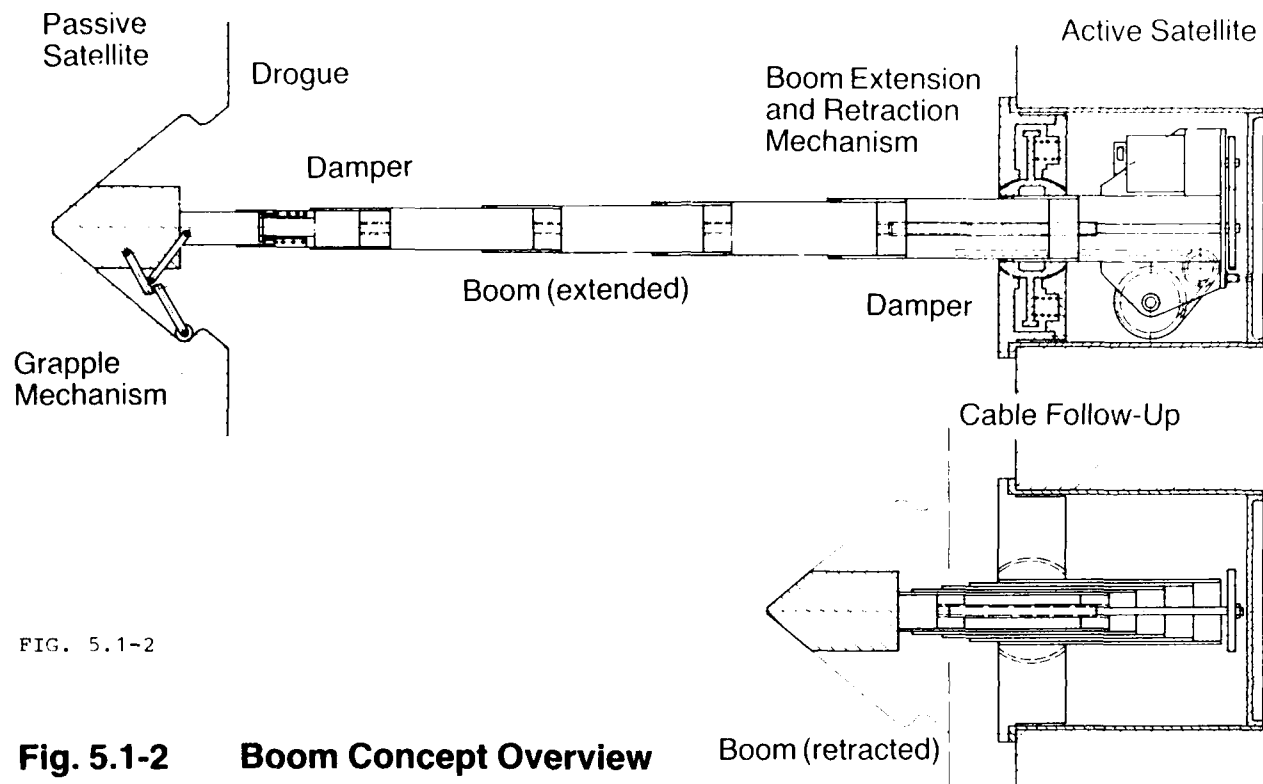


FIG. 5.1-2

Fig. 5.1-2 Boom Concept Overview

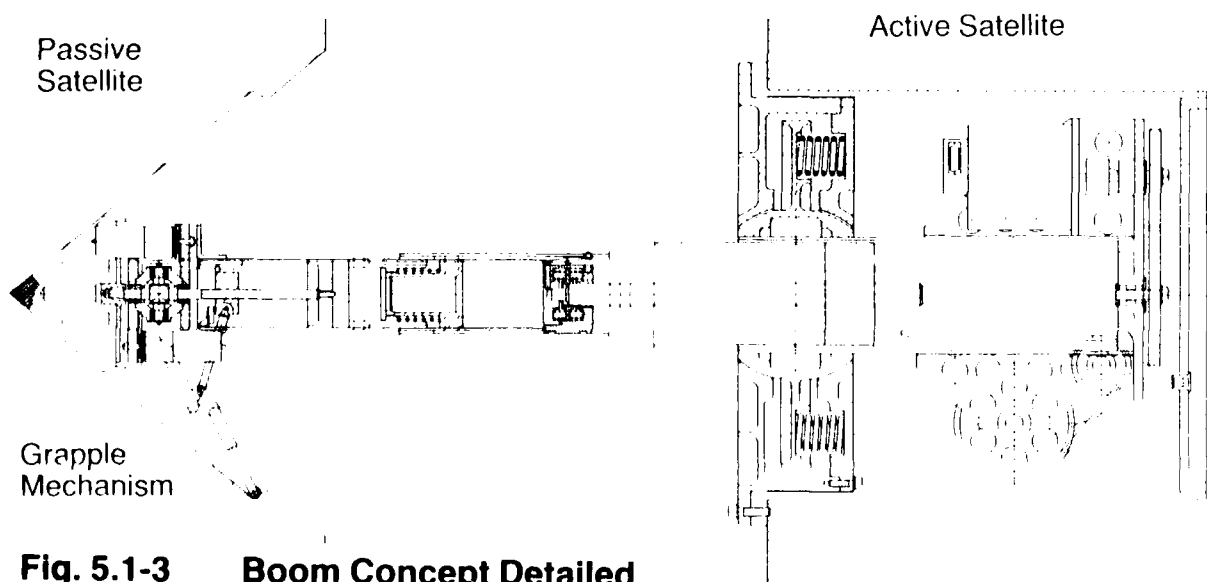


Fig. 5.1-3 Boom Concept Detailed

To make a trade-off between at least two designs and to evaluate to a better degree the probe/drogu concept, a claw-type grapple mechanism was designed. It was shown, that such a concept is feasible too and has its benefits.

5.2 Docking Mechanism for AOCS Controlled Closure

The Initial Separation Requirements for this case allow to skip the stand-off device as necessary element on the DMS controlled closure concepts. However it adds on the Active Satellite the following elements:

- special close-up sensors
- a derivative of the structural latches as used on the DMS controlled closure version.

This overall Docking Mechanism concept leads to the following General description for the AOCS controlled closure (see Table 5.2-1).

GENERAL DESCRIPTION		
ZERO IMPACT, NO-PROBE, NO-DROGUE SYSTEM		
FUNCTIONS	ACTIVE SATELLITE	PASSIVE SATELLITE
Guidance	Special close-up sensors deliver information to guide the Active Satellite within the gathering range of the grappling latches.	Cruciform stand-off device (no active parts)
Grappling Closure Latching	Combined in one mechanism (typically: latches) which does these functions subsequently. Damping is included within the mechanical and electro/mechanical units.	
Connection	Add-on features, independant of latches.	
Undocking	Opening of latches and separation velocity provided by AOCS.	
Emergency Provisions	- Pyrotechnics Devices - springs to guarantee a minimum separation velocity.	

TABLE 5.2-1: GENERAL DESCRIPTION OF AOCS CONTROLLED CLOSURE CONCEPT

Special attention within the AOCS controlled closure concept has to be paid for the close-up sensors. Besides the "normal" relative attitude data like relative range, range rate and angular offsets, the docking surfaces of the two spacecraft have to be considered. Narrowing two spacecraft to within 0.1 m, is only possible, if the two docking surfaces match to each other without touching. Taking into account the safety philosophy and some possible restrictions from the approach corridor during the final RVD phase, this leads you to a RVD sensor concept with:

- redundant, overlapping measuring ranges.

To give an example:

- medium range sensors: from 1000 m to 1 m
- near range sensors : from 5 m to 0.1 m
- close-up sensors : from 0.3 m to 0.05 m

The sensors specifically regarded for the AOCS controlled closure are the near range and the close-up sensor. As a design and selection guideline for these sensors the following should apply:

- sensors should be as simple as possible
- sensors should as far as possible rely on passive means on the passive satellite
- sensors shall withstand the service life without degradation or have the possibility of calibration, also by the sensors with overlapping range.

As a possible solution, the following concepts have emerged:

for the near range sensors: a system of position sensitive detectors on the Active Satellite, illuminated by LEDs on the Passive Satellite

for the close-up sensors : a system of proximity sensors, self illuminating by infrared phototransistors and receivers on the Active Satellite and a small reflective foil on the Passive Satellite.

5.3 Docking Mechanism Electronics

The Docking Mechanism Electronics (DME) has to be imbedded in the total RVD/AOCS-electronics concept which does not affect the possibility to design the DME in its own boxes for integration in the DMS. Major influence on the DME concept has the imposed safety philosophy. The requirements derived from there for the DME, which should be imposed at least for the RVD Electronics as well, lead to a DME concept based on:

- fault tolerant computing, by
- ultra-reliable computer system, with
- triple modular, dynamic redundancy, and a
- triple bus system

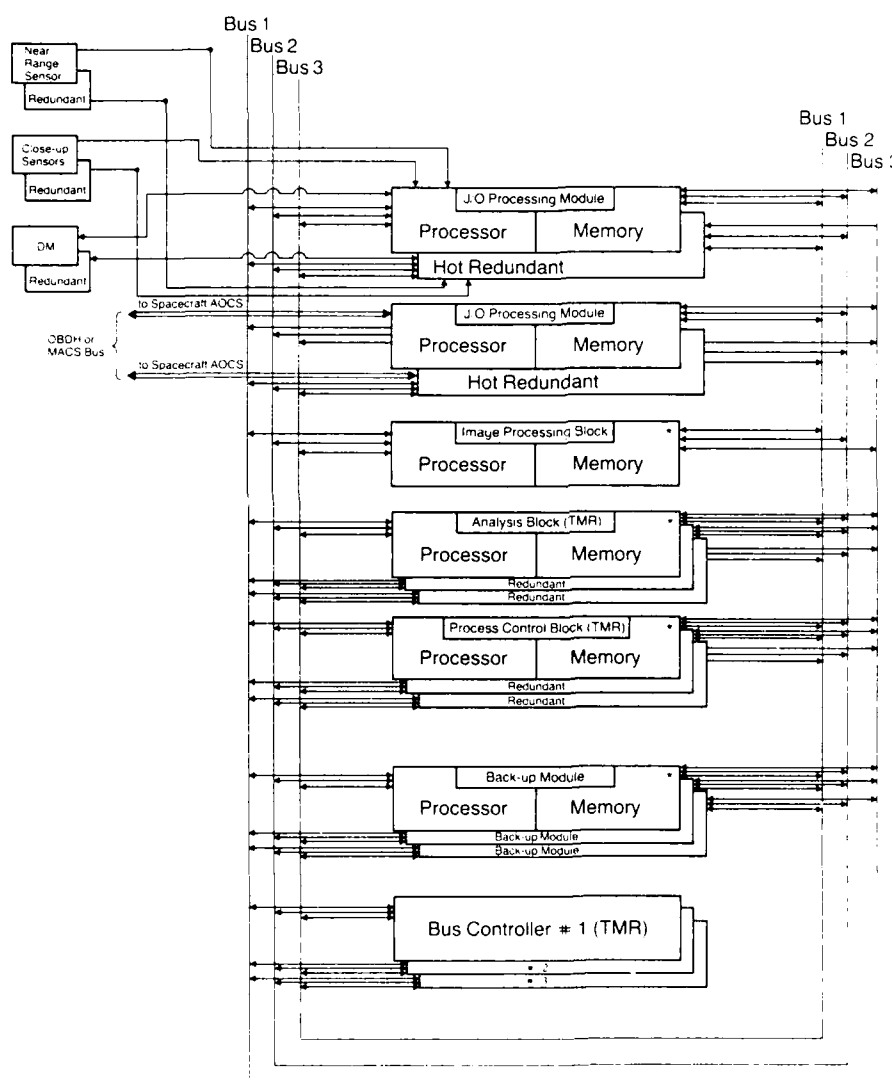
A general outline of this structure is given in Fig. 5.3-1 with the following major blocks: The Process Control Block (PCB) performs the overall control of the RVD procedure and coordinates major timing functions, processes the commands to and from the spacecraft, and makes decisions which RVD procedures are executed or aborted. So the PCB executes the normal docking sequence and handles exceptions and faults.

The RVD Sensing Block (RSB) handles the near range and close-up sensors and commands and controls the electro mechanical units in the Docking Mechanism.

The Image Processing Block (IPB) extracts pertinent information from the image data generated by the RVD sensors and converts the raw sensor data into image features. It also eliminates errors or takes information from the image.

The Analysis Block (AB) interprets the image features and determines the relative attitude of the two spacecrafts. As this is a very important role, the AB is fault tolerant.

The Back-Up Module (BUM) is a processor and memory area where any of the before mentioned processors may have access for reconfiguration in case of failure.



TMR: Triple Modular Redundancy
 *: Block Coupled by Dynamic Redundancy

Fig. 5.3-1 RVD-Data Electronics Concept for Fault Tolerant Computing by Ultra Reliable Computer System

The Docking Mechanism Power Electronics has to reflect the DMS safety philosophy as well. It features a dual independent main power bus for the nominal docking and undocking process and an emergency power bus to power the emergency undocking units within the DMS.

5.4 Connection Mechanism

The design concepts of the Connection Mechanism for satellite servicing obey in principle to the same general requirements as the Docking Mechanism itself does. From this point of view the following design guidelines are obvious:

- elements on the passive satellite have to be reusable after emergency undocking
- no pyrotechnics on passive satellite
- fixed connector plates on passive satellite
- minimum travel distance to mate both connector plates
- no structural load path through the connectors

With these guidelines in mind, there are still many possible solution for the Connection Mechanism, for example:

- single central or multiple peripheral actuators to move the connector plate
- quadratic or rectangular connector plate
- multiple or single connector actuation
- unidirectional or bidirectional actuation direction.

Designed in that way, the Connection Mechanism forms an integral, separately mountable subassembly within the Docking Mechanism Subsystem.

THE INFLUENCE OF SENSOR AND THRUSTER IMPERFECTIONS
ON THE ATTITUDE AND POSITION
OF A SPACECRAFT PERFORMING A RENDEZ-VOUS

BY

R. ANCILLOTTI and C. CASSI
AERITALIA - Space Systems Division
Via Servais, 125
10146 TORINO
Italy

LIST OF SYMBOLS

$\vec{g}(R)$	=	gravitational acceleration
\vec{a}_t	=	thrust acceleration on the target
\vec{a}_c	=	" " " the chaser
R_t, R_c	=	position of target and chaser center of mass
t	=	time
G	=	universal constant of gravitation
M	=	earth mass
ω_o	=	angular velocity of the earth
τ_o	=	phase angle at $t = 0$ when RV begins
θ, ϑ, ψ	=	Euler angles
$\dot{\theta}, \dot{\vartheta}, \dot{\psi}$	=	Euler angular rates
A, \dot{A}	=	right ascension and right ascension rate
L, \dot{L}	=	declination and declination rate
$\rho, \dot{\rho}$	=	range and range rate
$\theta, \dot{\theta}$	=	generical angle and angular rate
ϵ_θ	=	maximum measurement angular error
$A = N \epsilon_\theta$	=	maximum allowed angular error (phase plane)
$N \geq 4$	=	proper multiplier of the sensor error
\vec{C}	=	cartesian state vector
\vec{P}	=	polar state vector
$\Delta A, \Delta \dot{A}, \Delta L, \Delta \dot{L}, \Delta \rho, \Delta \dot{\rho}$	=	intrinsic error on the laser radar absolute value
$\Delta X, \Delta Y, \Delta Z, \Delta \dot{X}, \Delta \dot{Y}, \Delta \dot{Z}$	=	measurement error on the cartesian coordinates absolute value
\vec{A}	=	rotation matrix
$\vec{U}_c^{LOS}, \vec{U}_t^{LOS}$	=	line of sight direction in the chaser and target frame
$\vec{U}_c^{SUN}, \vec{U}_t^{SUN}$	=	sun direction in the chaser and target frame
γ_c, γ_t	=	LOS direction in-plane angle in the chaser and target frame
δ_c, δ_t	=	LOS direction out-of-plane angle in the chaser and target frame
β_c, β_t	=	sun direction out-of-plane angle in the chaser and target frame
F_{XN}, F_{YN}, F_{ZN}	=	nominal thrusts in the X, Y, Z directions
I_A	=	impulse for the attitude control
Δt	=	minimum actuation time
T_a	=	thrust level of the attitude engines
T_p	=	thrust level of the position engines
I	=	minimum impulse
M_p	=	propellant consumption
N_i	=	maximum number of impulses
\dot{S}	=	propellant consumption rate
V_c	=	nominal velocity in X and Z directions to eliminate lateral misalignments
ϵV	=	nominal measurement error on the lateral velocities \dot{X} and \dot{Z}

SUMMARY

The operational conditions of two spacecrafts performing a rendez-vous in geostationary orbit are given with the relative motion equation and the relevant control law. The concept of position, velocity and attitude measurement of the chaser vehicle with respect to the target one is given in order to allow the evaluation of the measurement errors.

Simulations show the effect of the errors sources on the flight path of the approaching vehicle and point out the necessity of a good matching between the sensor accuracy and the minimum impulse level of the actuation system.

1 - INTRODUCTION

The Rendez-vous manoeuvre results to be a very important phase in the assembling of large space structures by multiple launches of dockable modules. Telecommunication platforms require, in particular, Rendez-vous and docking manoeuvre in geostationary orbit.

The following operational conditions and system layout have been assumed:

The target is a rigid satellite waiting in GEO on station position for the chaser one.

It is equipped with the standard instrumentation for the three-axes stabilization plus a laser radar instrumentation. The chaser starts the RV manoeuvre when it falls into a 100 km neighborhood from the target and the mutual acquisition is established.

Launch, circularization and orbital drifting phases (to recover the angular displacement separating it on GTO from the target) are intended to be already completed.

Chaser is a rigid three axes stabilized spacecraft and its RV instrumentation consists of:

- i) An optical telescope placed in front of it in order to receive the light emitted by the target, necessary for relative attitude measurement;
- ii) A certain number of corner reflectors in order to reflect back the signal toward the target;
- iii) A sun sensor for attitude measurement.

A S-band link between target and chaser is provided in order to allow the exchange of informations and actuation signals.

The target spacecraft (T. S/C) is assumed to be completely controlled either in attitude and in position. The chaser attitude and position are controlled by an all thrusters actuation system. Figure 1 gives the sensor and thruster layout of the chaser spacecraft (C. S/C).

2 - RENDEZ-VOUS MATHEMATICAL MODEL

Considering the T. S/C and C. S/C centers of mass (see fig. 2) and referring to an inertial frame centered in the earth, the vectorial equation of the relative motion (C. S/C with respect to the T. S/C) is:

$$\frac{D^2 \vec{s}}{Dt^2} = \vec{a}_c - \vec{a}_t + \vec{g}(\vec{R}_c) - \vec{g}(\vec{R}_t)$$

neglecting: . earth oblateness
 . atmospheric drag perturbation
 . luni-solar perturbations

and assuming $S \ll R$ one can write: (Ref. 1)

$$\frac{D^2 \vec{s}}{Dt^2} = \vec{a}_c - \vec{a}_t - \frac{GM}{R_t^2} \left[\frac{\vec{s}}{R_t} - 3 \frac{\vec{R}_t (\vec{s} \cdot \vec{R}_t)}{R_t^3} \right]$$

Assuming a target fixed ref. frame centred in its GC, which translates maintaining its axes parallel to the earth fixed inertial ref. frame, the following three equations are obtained assuming $\vec{a}_t = 0$ (i.e. only the chaser performs active RV):

$$\ddot{x} = a_x - \omega_o^2 \left\{ x \left[1 - 3 \cos^2(\omega_o t + \tau_o) \right] - 3/2 y \sin(2\omega_o t + 2\tau_o) \right\}$$

$$\ddot{y} = a_y - \omega_o^2 \left\{ -3/2 x \sin(2\omega_o t + 2\tau_o) + y \left[1 - 3 \sin^2(\omega_o t + \tau_o) \right] \right\}$$

$$\ddot{z} = a_z - \omega_o^2 z$$

Assuming a chaser fixed reference frame (x_B, y_B, z_B) as in fig. 3, the attitude motion of the chaser in this frame is described by the three Euler equations: (Ref. 2)

$$\begin{aligned}\dot{P} &= T_{XB}/A - QR(C-B)/A \\ \dot{Q} &= T_{YB}/B - PR(A-C)/B \\ \dot{R} &= T_{ZB}/C - PQ(B-A)/C\end{aligned}$$

Referring now the angular motion to a (x_1, y_1, z_1) frame, parallel to the (x, y, z) one centered in the target, it is possible to write:

$$\begin{aligned}\dot{\theta} &= P + Q \sin \theta \tan \Theta + R \cos \theta \tan \Theta \\ \dot{\phi} &= Q \cos \theta - R \sin \theta \\ \dot{\psi} &= (Q \sin \theta + R \cos \theta) / \cos \Theta\end{aligned}$$

3 - POSITION AND VELOCITY CONTROL LAWS

Two complementary strategies have been adopted:

- velocity control at long range from the target ("collision course")
- position-velocity control at short range performing the line of sight-docking axis alignment

Strategy (a) is accomplished following the phase diagram of fig. 4 showing the dependence of the range rate $\dot{\rho}$ on the range ρ .

This type of control forces the velocity of the chaser to be completely radial with respect to the given reference frame centered on the target.

$$\begin{aligned}\dot{x}_N &= \dot{\rho}_N \cos L \cos A \\ \dot{y}_N &= \dot{\rho}_N \cos L \sin A \\ \dot{z}_N &= \dot{\rho}_N \sin L\end{aligned}$$

being $A = L = 0$ for radial motion (See fig. 5), and $\dot{\rho}_N$ the nominal range rate.

Fig. 6 shows the block diagram of the procedures to be performed by on-board computer in order to control this approaching phase.

Strategy (b) is accomplished controlling (x, \dot{x}) and (z, \dot{z}) according to the ON-OFF concept of fig. 7 and (y, \dot{y}) following the phase plane diagram of fig. 4. An analogous block diagram of the on-board computer procedures is shown in fig. 8.

4 - CHASER ATTITUDE CONTROL LAWS

The chaser vehicle is three axis stabilized via an ON-OFF control strategy independently operating on each axis as fig. 9 shows. The $(\theta, \dot{\theta})$ status is maintained within the plane strip limited by the two straight lines:

$$\theta + \epsilon \dot{\theta} = A \quad \text{and} \quad \theta + \epsilon \dot{\theta} = -A \quad \text{where} \quad A = N \epsilon_{\theta} \quad N \geq 4$$

The body reference frame (x_B, y_B, z_B) of the chaser is always maintained parallel to the frame centered in the target (see fig. 10).

5 - CHASER POSITION MEASUREMENT CONCEPT

Position and velocity of the chaser vehicle with respect to the target are directly measured on the latter, collecting by the optical telescope the light reflected back from the corner reflectors of the chaser.

These measures will be obtained in a polar form:

$$\vec{P} = (P_1, P_2, P_3, P_4, P_5, P_6) = (\rho, A, L, \dot{\rho}, \dot{A}, \dot{L})$$

and, if necessary, will be transformed in cartesian form, using the well known transformations to give:

$$\vec{C} = (C_1, C_2, C_3, C_4, C_5, C_6) = (x, y, z, \dot{x}, \dot{y}, \dot{z})$$

With the aim of comparing two types of laser-radar (see ref. 3) a first evaluation of the absolute error on the cartesian state vector \underline{C} as a function of the measurement errors on the polar state vector \underline{P} , has been performed introducing the error propagation formula at first order:

$$\Delta C_J = \sum_{i=1}^6 \left| \frac{\partial C_J}{\partial P_i} \right| \Delta P_i \quad J=1, \dots, 6$$

Hereafter the results are shown for two different laser radar devices (i), (ii) along with the operational conditions chosen for the numerical evaluations.

The following angular positions and rates for the chaser are chosen in all the \underline{P} , $\dot{\underline{P}}$ cases:

$$A = -90^\circ = -1.5708 \text{ rad}$$

$$\dot{A} = 7.29 \cdot 10^{-5} \text{ rad/s} \quad (\text{earth rotation rate})$$

$$L = 0.01 \text{ rad}$$

$$\dot{L} = .0 \text{ rad/s}$$

i) sensor errors

$$3.000 \text{ m} < \rho < 120.000 \text{ m} \text{ (long range)}$$

$$\Delta \rho = 0.5 \%$$

$$\Delta \dot{\rho} = 0.2 \%$$

$$\Delta A = \Delta L = 0.1^\circ = 0.005 \text{ rad}$$

$$\Delta \dot{A} = \Delta \dot{L} = 0.5 \cdot 10^{-3} \text{ rad/s}$$

$$0 < \rho < 3000 \text{ m} \text{ (short range)}$$

$$\Delta \rho = 0.1 \text{ m}$$

$$\Delta \dot{\rho} = 0.03 \text{ m/s}$$

$$\Delta A = \Delta L = 0.005 \text{ rad}$$

$$\Delta \dot{A} = \Delta \dot{L} = 0.5 \cdot 10^{-4} \text{ rad/s}$$

The results are shown in table I.

	$\rho = 50000 \text{ M}$ $\dot{\rho} = 3.5 \text{ M/SEC}$	$\rho = 10000 \text{ M}$ $\dot{\rho} = 1.7663 \text{ M/SEC}$	$\rho = 1000$ $\dot{\rho} = 0.206 \text{ M/SEC}$	$\rho = 100$ $\dot{\rho} = 0.05 \text{ M/SEC}$
ΔX	84.9957 M	16.9991 M	1.6999 M	0.1699 M
ΔY	250.8375 M	50.1675 M	0.1169 M	0.1017 M
ΔZ	87.4957 M	17.4491 M	1.7010 M	0.1710 M
$\Delta \dot{X}$	25.0230 M/SEC	5.0064 M/SEC	0.5003 M/SEC	0.0501 M/SEC
$\Delta \dot{Y}$	0.0230 M/SEC	0.0613 M/SEC	0.0351 M/SEC	0.0305 M/SEC
$\Delta \dot{Z}$	25.0048 M/SEC	5.0028 M/SEC	0.5006 M/SEC	0.0504 M/SEC

ii) sensor errors

$$\Delta \rho = 0.1 \text{ m}$$

$$\Delta \dot{\rho} = 10^{-5} \text{ m/s}$$

$$\Delta A = \Delta L = 10^{-4} \text{ rad}$$

$$\Delta \dot{A} = \Delta \dot{L} = 10^{-5} \text{ rad/s}$$

The results are shown in table II.

	$\rho = 50000 \text{ M}$	$\rho = 10000 \text{ M}$
ΔX	4.9997 M	0.9999 M
ΔY	0.1499 M	0.1099 M
ΔZ	5.0007 M	1.0009 M
$\Delta \dot{X}$	0.5003 M/s	0.1002 M/s
$\Delta \dot{Y}$	5.3779×10^{-3} M/s	1.0846×10^{-3} M/s
$\Delta \dot{Z}$	0.5003 M/s	0.1002 M/s

6 - CHASER ATTITUDE MEASUREMENT CONCEPT

The attitude of the chaser is obtained measuring the line-of sight (\vec{U}_{LOS}) and the sun direction vector (\vec{U}_{SUN}) either in the target and in the chaser reference frame.

Being \vec{A} the linearised (3, 2, 1) rotation matrix from target to chaser reference frame and ψ, θ, ϕ the Euler angles, it is possible to write:

$$\vec{A} = \begin{bmatrix} 1 & \psi & -\theta \\ -\psi & 1 & \phi \\ \theta & -\phi & 1 \end{bmatrix}$$

$$\vec{U}_{LOS}^C = \vec{A} \vec{U}_{LOS}^T \quad (1)$$

$$\vec{U}_{SUN}^C = \vec{A} \vec{U}_{SUN}^T \quad (2)$$

referring to fig. 11 it is:

$$\vec{U}_{LOS}^T = (U_X^T, U_Y^T, U_Z^T) = (\cos \delta_T \cos \delta_T, \cos \delta_T \sin \delta_T, \sin \delta_T)$$

$$\vec{U}_{LOS}^C = (U_X^C, U_Y^C, U_Z^C) = (\cos \delta_C \cos \delta_C, \cos \delta_C \sin \delta_C, \sin \delta_C)$$

$$\vec{U}_{SUN}^T = (U_X^T, U_Y^T, U_Z^T) = (\cos \beta_T \cos \alpha_T, \cos \beta_T \sin \alpha_T, \sin \beta_T)$$

$$\vec{U}_{SUN}^C = (U_X^C, U_Y^C, U_Z^C) = (\cos \beta_C \cos \alpha_C, \cos \beta_C \sin \alpha_C, \sin \beta_C)$$

Combining the first two equations (1) and the last equation (2) we have the following linear system in the variables ψ, θ, ϕ :

$$\psi \cos \delta_C \sin \delta_C - \theta \sin \delta_C = \cos \delta_C \cos \delta_C - \cos \delta_T \cos \delta_T$$

$$\psi \cos \delta_C \cos \delta_C - \phi \sin \delta_C = \cos \delta_T \sin \delta_T - \cos \delta_C \sin \delta_C$$

$$\theta \cos \beta_C \cos \alpha_C - \phi \cos \beta_C \sin \alpha_C = \sin \beta_C - \sin \beta_T$$

The system has been resolved by the Cramer method.

Let D be the coefficients determinant:

$$D = \cos \delta_C \sin \delta_C \cos \beta_C (\sin \delta_C \cos \alpha_C - \cos \delta_C \sin \alpha_C)$$

Then the Euler angles expression as function of the angles computed by the sensors are:

$$\psi = \left[(\sin \beta_C - \sin \beta_T) \sin^2 \delta_C - \sin \delta_C \cos \beta_C \sin \alpha_C (\cos \delta_T \sin \delta_T - \cos \delta_C \sin \delta_C) + \sin \delta_C \cos \beta_C \cos \alpha_C (\cos \delta_C \cos \delta_C - \cos \delta_T \cos \delta_T) \right] D^{-1}$$

$$\theta = \left[\cos \delta_C \sin \delta_C \cos \beta_C \sin \alpha_C (\cos \delta_C \sin \delta_C - \cos \delta_T \sin \delta_T) + \cos \delta_C \cos \delta_C \cos \beta_C \sin \alpha_C (\cos \delta_C \cos \delta_C - \cos \delta_T \cos \delta_T) + \cos \delta_C \sin \delta_C \sin \delta_C (\sin \beta_C - \sin \beta_T) \right] D^{-1}$$

$$\phi = \left[\cos \delta_C \cos \delta_C \cos \beta_C \cos \alpha_C (\cos \delta_C \cos \delta_C - \cos \delta_T \cos \delta_T) + \cos \beta_C \cos \alpha_C \cos \delta_C \sin \delta_C (\cos \delta_C \sin \delta_C - \cos \delta_T \sin \delta_T) + \sin \delta_C \cos \delta_C \cos \delta_C (\sin \beta_C - \sin \beta_T) \right] D^{-1}$$

A first evaluation of $\Delta\psi, \Delta\theta, \Delta\phi$ have been made applying again the error propagation formula.

Numerical examples show that variation of ψ, θ, ϕ are very light in the L.O.S. oscillation range during RV.

Being $\Delta\alpha_{SS} = .1^\circ$ for the sun sensor measured angles

and $\Delta\alpha_{LR} = .006^\circ$ for the laser-radar measured angles, the errors on Euler angles result to be:

$$\Delta\phi = \Delta\psi = .031^\circ$$

$$\Delta\theta = .11^\circ$$

7 - THRUSTERS MISALIGNMENT

Thrusters mechanical misalignment is taken into account in fig. 12 where θ is a chaser corner and F_{XN}, F_{YN}, F_{ZN} represent the forces due to the three thrusters mounted on that corner.

Everyone of these devices is misaligned with respect to its nominal direction by angles γ and γ' as

figure shows. Because the standard misalignment is about 0.1%, the perturbing forces are neglectable being the 1 % of the nominal ones.

8 - NUMERICAL SIMULATIONS

Two different simulations have been done starting from the common initial conditions of Table III. The two simulations represent the behaviour of the approaching vehicle equipped with the different measurement systems. (Table IV: coarser sensor errors, Table V: more accurate sensor errors).

TABLE III : INITIAL CONDITIONS

$X = 5 \text{ m}$	$\dot{X} = 0.1 \text{ m/s}$	$\emptyset = \theta = \psi = 0.001 \text{ rad}$
$Y = -60 \text{ m}$	$\dot{Y} = 0.95 \text{ m/s}$	$\dot{\emptyset} = \dot{\theta} = \dot{\psi} = 0.0001 \text{ rad/s}$
$Z = 5 \text{ m}$	$\dot{Z} = 0.1 \text{ m/s}$	

TABLE IV : COARSER SENSOR ERRORS

$E\emptyset = 0.1 \text{ m}$	$EA = EL = 1.75 \cdot 10^{-3} \text{ rad} = 0.1^\circ$
$E\dot{\emptyset} = .03 \text{ m/s}$	$E\dot{A} = E\dot{L} = 5 \cdot 10^{-5} \text{ rad/s}$
$E\emptyset = E\theta = E\psi = .87 \cdot 10^{-2} \text{ rad}$	
$E\dot{\emptyset} = E\dot{\theta} = E\dot{\psi} = .87 \cdot 10^{-3} \text{ rad/s}$	

TABLE V : MORE ACCURATE SENSOR ERRORS

$E\emptyset = 0.1 \text{ m}$	$EA = EL = .1 \cdot 10^{-3} \text{ rad}$
$E\dot{\emptyset} = 0.1 \cdot 10^{-4} \text{ m/s}$	$E\dot{A} = E\dot{L} = .1 \cdot 10^{-4} \text{ rad/s}$
$E\emptyset = E\theta = E\psi = .524 \cdot 10^{-4} \text{ rad}$	
$E\dot{\emptyset} = E\dot{\theta} = E\dot{\psi} = .524 \cdot 10^{-4} \text{ rad/s}$	

Tables VI and VII give the final conditions for the two cases.

TABLE VI : FINAL CONDITIONS : 1st ERROR SET

$x = -1.21 \cdot 10^{-1} \text{ m}$	$\dot{x} = 0.0015 \text{ m/s}$
$y = -4.2 \text{ m}$	$\dot{y} = -0.004 \text{ m/s}$
$z = 9.33 \cdot 10^{-2} \text{ m}$	$\dot{z} = 0.00345 \text{ m/s}$

$\emptyset = 0.0525 \text{ rad} (= 1.44^\circ) < 0.04 \text{ rad} (= 2.29^\circ)$	$\dot{\emptyset} = 0.00076 \text{ rad/s} < 0.00128 \text{ rad/s}$
$\theta = 0.041 \text{ rad} (= 2.35^\circ) < 0.043 \text{ rad} (= 2.46^\circ)$	$\dot{\theta} = 0.000468 \text{ rad/s} < 0.00115 \text{ rad/s}$
$\psi = 0.0108 \text{ rad} (= 0.619^\circ) < 0.0125 \text{ rad} (= 0.71^\circ)$	$\dot{\psi} = 0.000191 \text{ rad/s} < 0.0007 \text{ rad/s}$

Attitude propellant consumption = 0.0053 KG
 Position propellant consumption = 1.584 KG
 Total propellant consumption = 1.6 KG

TABLE VII : FINAL CONDITIONS : 2nd ERROR SET

$x = 0.001 \text{ m}$	$\dot{x} = 0.00067 \text{ m/s}$
$y = -4.4 \text{ m}$	$\dot{y} = 0.00015 \text{ m/s}$
$z = 0.001 \text{ m}$	$\dot{z} = 0.00067 \text{ m/s}$

$\emptyset = -0.001 \text{ rad} (= -0.057^\circ) < -0.005 \text{ rad} (= -0.28^\circ)$	$\dot{\emptyset} < 0.0002 \text{ rad/s}$
$\theta = -0.01 \text{ rad} (= -0.57^\circ) < -0.015 \text{ rad} (= -0.86^\circ)$	$\dot{\theta} < 0.0002 \text{ rad/s}$
$\psi = -0.002 \text{ rad} (= -0.11^\circ) < -0.003 \text{ rad} (= -0.17^\circ)$	$\dot{\psi} < 0.0002 \text{ rad/s}$

Attitude propellant consumption = 0.813 KG
 Position propellant consumption = 1.325 KG
 Total propellant consumption = 2.138 KG

In table VI and VII the maximum values achieved during the simulations are given too.

The second case shows the following characteristics with respect to the first:

- i) better alignment
- ii) lower angular rates
- iii) lower lateral displacements and velocities
- iv) very higher propellant consumption for attitude control
- v) slightly lower propellant consumption for position and velocity control.

The main result of the simulations is that a higher attitude thrusters impulse level leads to a higher propellant consumption if the more accurate measurement set is maintained and if more stringent attitude control is applied.

Owing to this reason other two cases have been simulated maintaining the more accurate sensor system and control laws, but varying the impulse level:

I_a	Δt	T_a	T_p
0.2	0.2	1	10
0.02	0.2	0.1	10

Fig.s 13 and 14 report θ Euler angle and its angular rate $\dot{\theta}$ for $I = 0.2$ Ns. Fig.s 15 and 16 show the same parameters for $T = 0.02$ Ns.

The propellant consumption is reported in table VIII.

TABLE VIII	I Ns	M_p (g/s)	N_i	S (g/s)
	2	813	162	7.39
	0.2	5.3	16	4.82×10^{-2}
	0.02	5.32	147	4.83×10^{-2}

The results show a drop of it to a minimum of 5.3 g for 110 s of flight but an increase of the maximum number of impulses. The optimal condition is obtained for $I = 0.1$ Ns with only 16 impulses. Rendezvous time constraint suggests to increase the thrust level of the position control engines. For this reason we investigate the case where $T_p = 20$ N, $T_a = 1$ N and $t = 0.2$ s.

It has been observed:

- i) A doubling of the disturbing torque which leads to an increase of the attitude control frequency and of the relevant propellant consumption.
- ii) An increase of the propellant consumption of the thrusters devoted to the position and velocity control, due to the improper matching between the minimum control impulse and the narrow band adopted on the phase plane.

Fig. 17 represents the in-plane flight path. The out-of-plane path is quite similar due to the same initial conditions assumed for the simulation.

Fig. 18 shows the velocity profile, \dot{y} as a function of the time. Fig.s 19-20 evidence its oscillations in detail.

It can be observed that the position-velocity control frequency is increased. The control is not able to maintain the chaser within the ON-OFF boundaries but, due to the minimum impulse level too high, drives the (X, X) point in the phase plane, alternatively from a side to the other, out of the ON-OFF boundary.

The control is always working with a sensible increase of the propellant consumption, being $m_p = 2.623$ kg. The propellant for the attitude control is increased to $m_{pa} = 10.5$ g. Fig. 21 shows as an example the increased control frequency $V = 32$ pulses in 140 seconds for the attitude angle ψ .

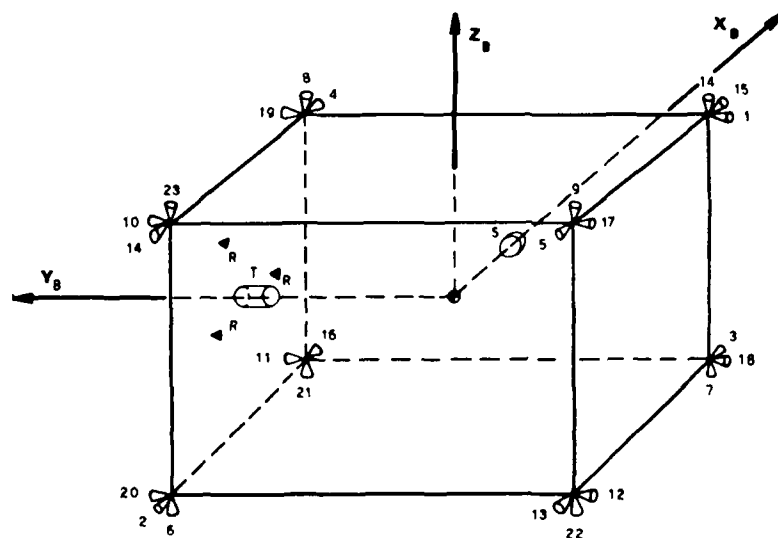
9 - CONCLUSIONS

This paper points out some problems inherent to the RV maneuver. It evidences that the measurement errors depend on the chaser and target status, so that the intrinsic sensor errors cannot be considered the maximum ones. Whilst the attitude measure error shows a light variability, instead the position and velocity ones are strongly variable. Simulations show the importance of a good matching between the sensor accuracy and the minimum impulse level of the actuation system in order to avoid a continuous oscillation in attitude of the spacecraft and an unacceptable propellant consumption.

The simulations presented in this paper are not exhaustive with all the disturbance effects during a RV manoeuvre. Nevertheless parametric nature of the developed software allows the introduction of them.

REFERENCES

- (1) Raymond W. WOLVERTON - "Flight performance handbook for orbital operations". John Wiley & Sons, 1967
- (2) Bernard ETKIN - "Dynamic of flight" . John Wiley & Sons, Inc 6th printing, 1967
- (3) Charles L. WYMAN - "Test performance of an experimental laser radar for rendez-vous and docking". J. SPACECRAFT VOL. 5, No. 4.



T: TELESCOPE, S: SUN SENSOR, R: CORNER REFLECTORS

Fig.1 Layout of the chaser spacecraft

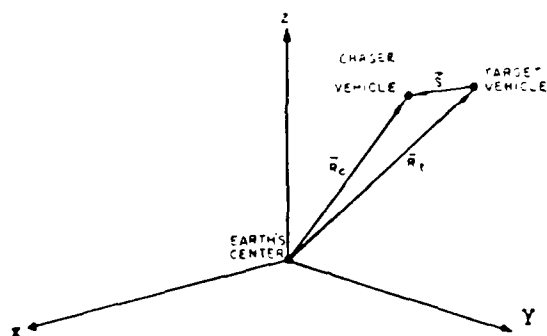


Fig.2 Inertial reference frame

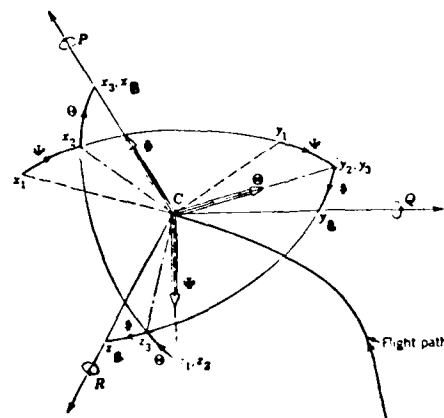


Fig.3 Chaser fixed reference frame (x_B, y_B, z_B)

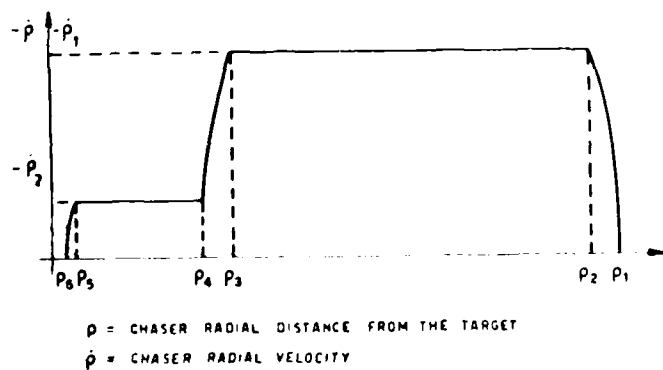


Fig.4 Phase plane ($\rho, \dot{\rho}$) diagram
position control

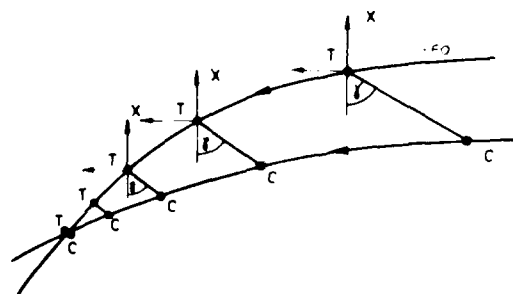


Fig.5 Collision course
T: target C: chaser

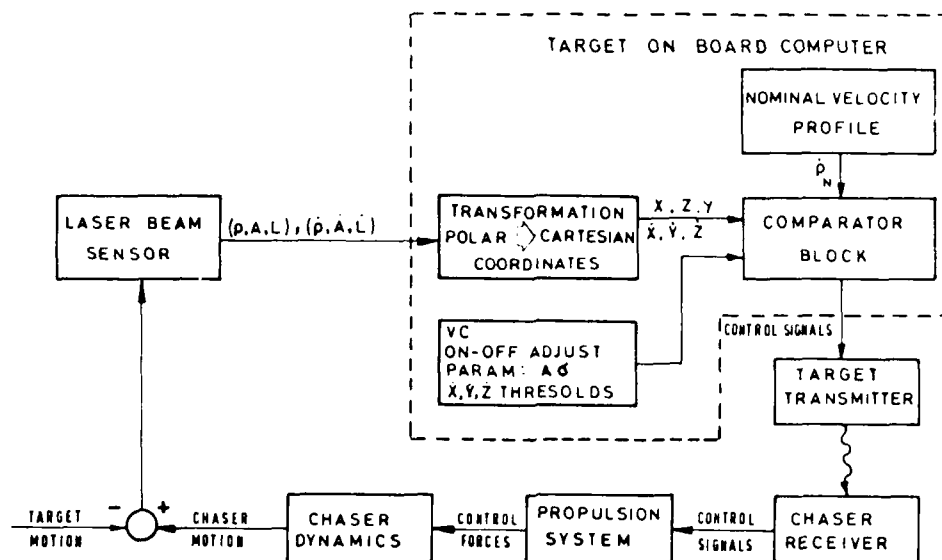


Fig.6 On board procedures block diagram: phase a)

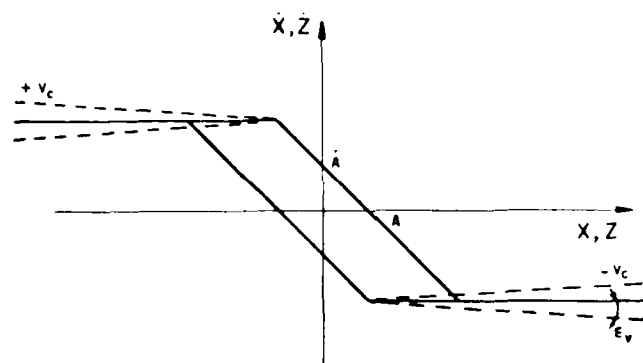


Fig.7 Phase plane of the on-off strategy

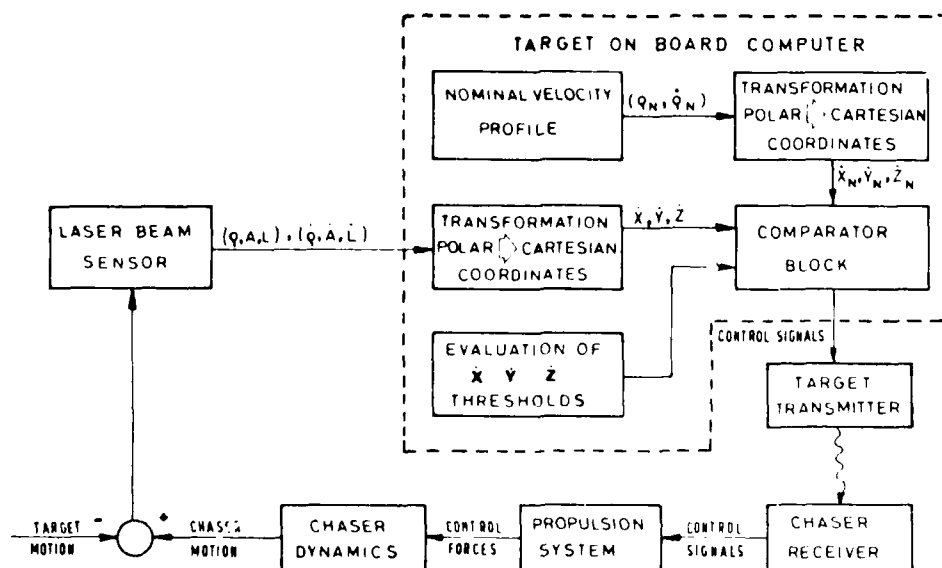


Fig.8 On board procedures block diagram: phase b)

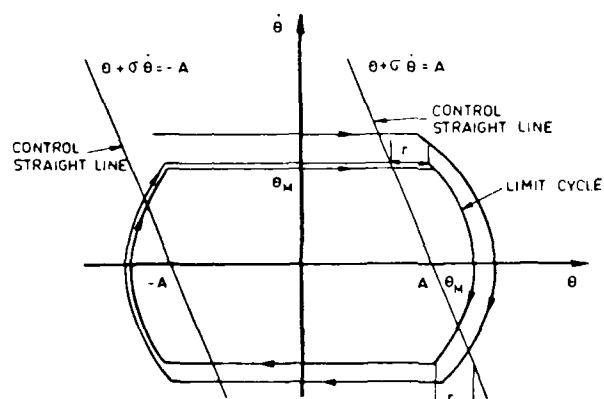
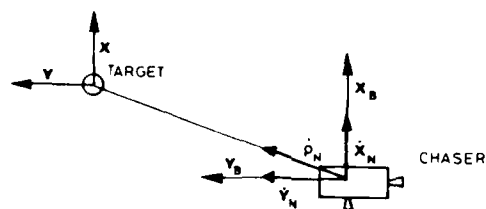
Fig. 9 On-off phase diagram $(\theta, \dot{\theta})$ 

Fig. 10 Target and chaser body frames relative position

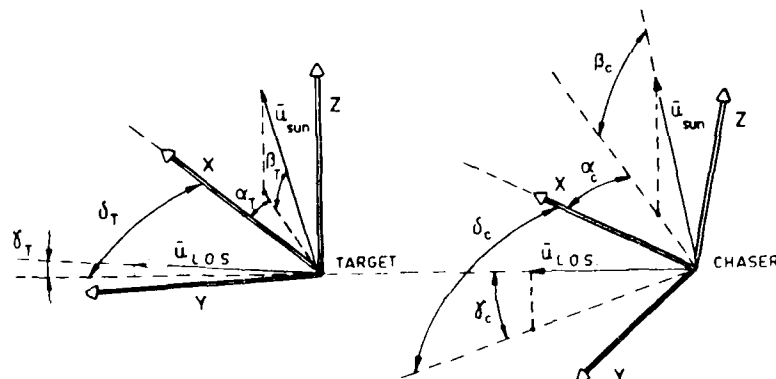


Fig. 11 Sun direction and L.O.S direction in both target and chaser reference frame

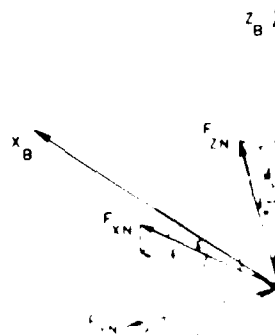


Fig. 12 Thruster misalignment

AD-A141 969

CONFERENCE PROCEEDINGS ON GUIDANCE AND CONTROL
TECHNIQUES FOR ADVANCED SP. (U) ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY. JAN 84

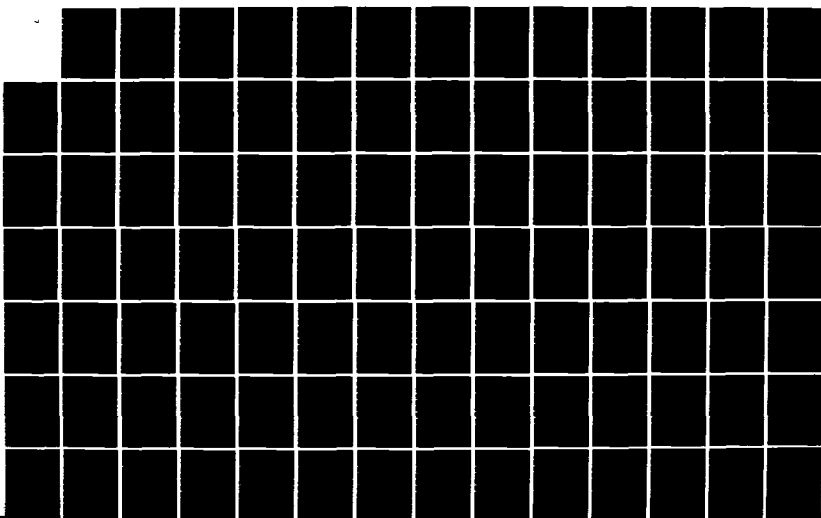
3/5

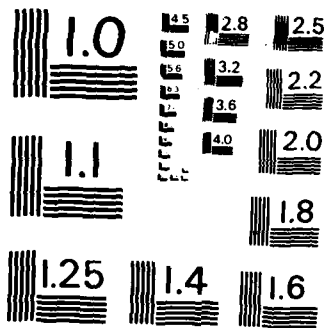
UNCLASSIFIED

AGARD-CP-358

F/G 22/3

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

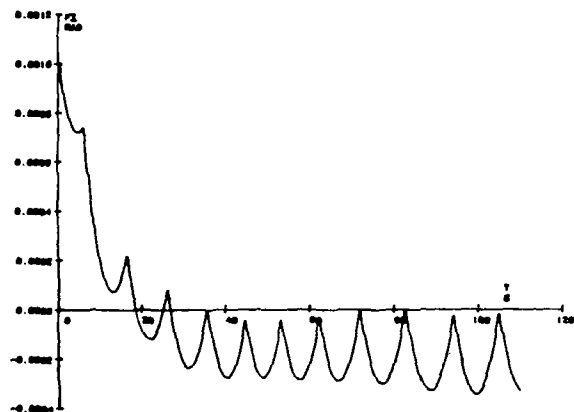


Fig. 13 Behaviour of the ϕ Euler angle during the time (0.2Ns impulse level)

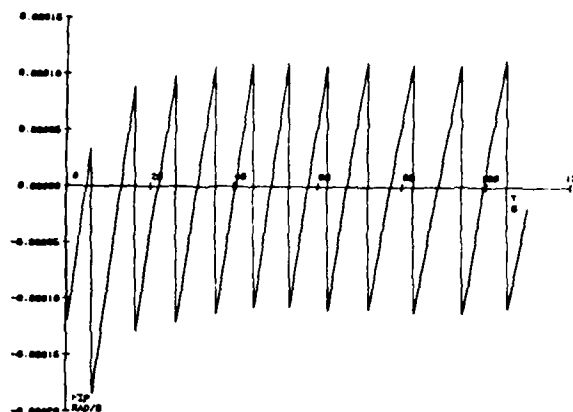


Fig. 14 The ϕ evolution during the time (0.2 Ns thruster impulse level)

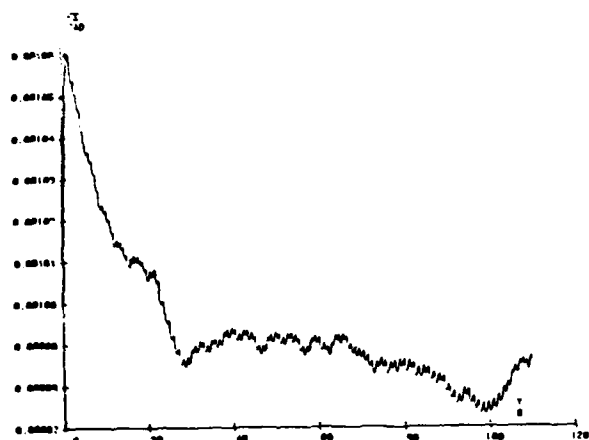


Fig. 15 ϕ evolution during the time (0.02Ns thruster impulse level)

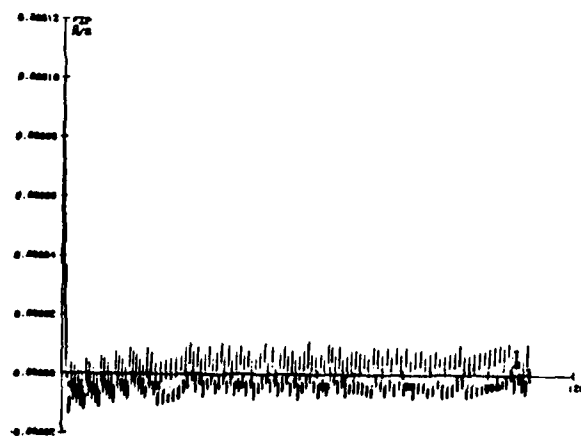


Fig. 16 ϕ evolution during the time (0.02Ns thruster impulse level)

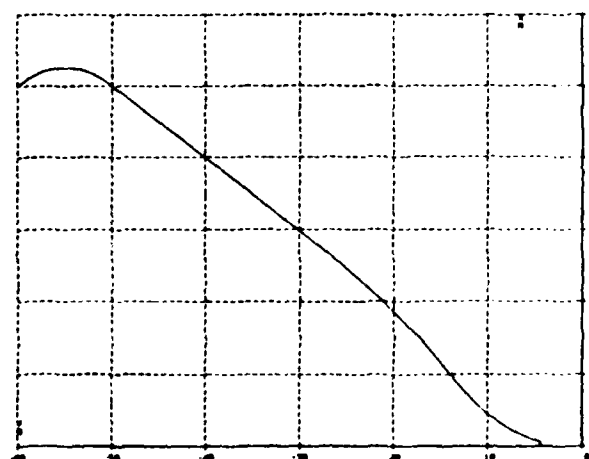


Fig. 17 In plane flight path

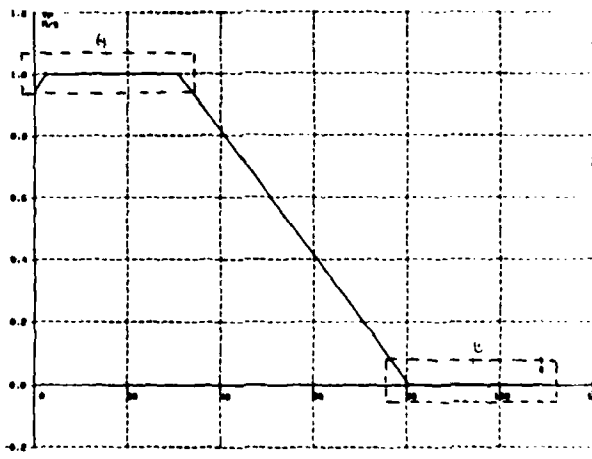


Fig. 18 Velocity profile \dot{y} function of the time

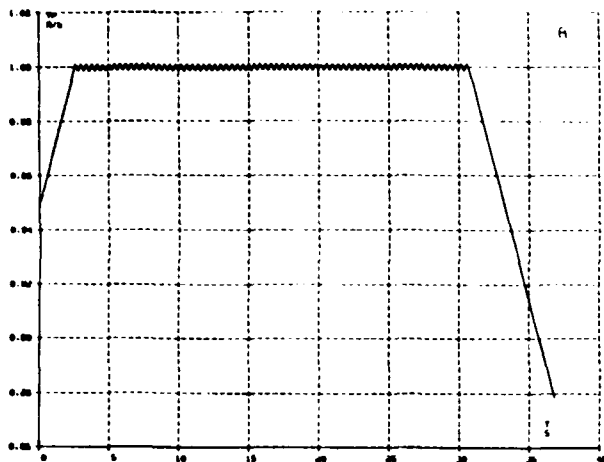


Fig.19 Detail of the y oscillation
during the time (particular a)

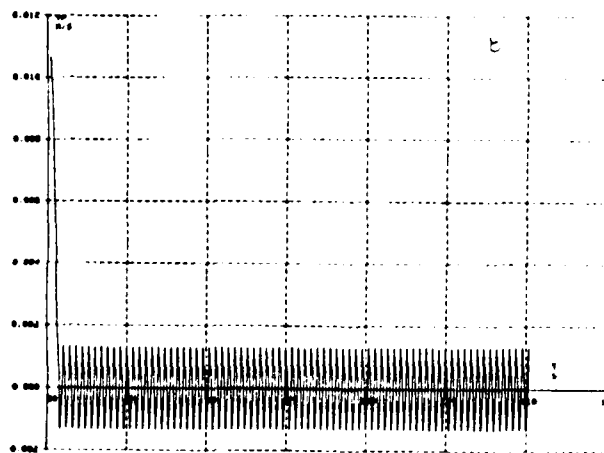


Fig.20 Detail of the y oscillation
during the time(particular b)

7

AD P 003 392

CONTROL ASPECTS AS ELABORATED IN SPACE RENDEZVOUS SIMULATIONS

by

P. Natenbruk and D. Rangnitt
ERNO Raumfahrttechnik GmbH
Huenefeldstr. 1-5
2800 Bremen 1
West Germany

1. INTRODUCTION

With the dawn of the operational phase of the American Space Shuttle System, as well as of the European Ariane Satellite Launcher, the search for the technological needs of tomorrow, in space, has been intensified.

Thus, for the last few years, the European Space Agency (ESA) has been sponsoring an increasing system and technology study effort related to in-orbit operations, covering – in particular – the performance of Rendezvous and Docking between two spacecraft. These two spacecraft, the one actively doing the rendezvous, the other passively waiting, we hereafter term the chaser and its target, respectively.

Unlike the US missions, RVD here is supposed to be done without any man in space. The chaser does its operations to some extent automatically, having on board the degree of autonomy needed.

Otherwise, it is steered from the ground by real-time or time tagged commands depending upon the contact opportunities.

2. STUDIES APPLIED

To set the stage, let us begin with the scope of subjects to which the reported programme was applied.

Figure 1 compiles the main Rendezvous and Docking (RVD) related missions, studied by Industry for ESA. Covered are flights on

- | | | |
|---|---|------------|
| – circular low earth orbits | – | RVD in LEO |
| – the elliptical geostationary transfer orbit | – | RVD in GTO |
| – the circular geostationary orbit itself | – | RVD in GEO |

The objectives of the missions were assumed to be different, such as

- flight demonstration of RCD capabilities in LEO
- or assembly of a communication platform in GEO

The launcher involved as earth-to-orbit transport vehicle were *either* the

- Shuttle for LEO
- Ariane for GTO and GEO

Also the spacecraft participating in the rendezvous process

- the target and the chaser were assumed to be different per studied mission.

Of major relevance for the simulation programme were the different orbits, the requirements and constraints due to the mission specific objective, and the spacecraft characteristics.

Let us recall here that in general we divide the Rendezvous (RV) process into the following three phases:

- The Guided Phase – during which the chaser is fully steered from the ground to an acquisition range from the target,
 - The Homing Phase – during which the chaser aims at an arrival in the close vicinity of the target,
- and
- Final Approach – when final arrival conditions, e.g. for docking, are achieved.

The first phase is handled with conventional techniques and therefore is not treated here. The same applies to homing – provided it is again guided from the ground. Of prime interest is – hence – final approach (FA).

3. THE PROGRAMME STRUCTURE

Although a complete RVD simulation programme must incorporate guidance of the chaser motion along its flight path as well as attitude control of both S/C, modelling of the sensors and thrusters inclusive of signal and command error modelling, the programme presented here is restricted to the guided and non-guided motion of the chaser relative to the target without attitude control and error modelling. Thus, the S/C is considered to be a mass point, with its thrusters ideally oriented.

A brief description of the structure of the discussed FA simulation programme is given in Figure 2.

Input data, such as the orbital data or the initial relative state, are fed into the programme managing part, which activates numerical integration routines.

The latter incorporate differential equations of the chaser's motion relative to the target, as well as those describing the absolute motion of the target itself. Thus the programme simulates the navigational use of relative measurements, as could be done from on board, one of the spacecraft having direct viewing or radio frequency contact with the other one.

The routines used are built-up according to the different control modes which, in turn, make up the desired rendezvous strategy and are steered by the appropriate control parameters included in the input data.

Such control modes, per strategy, are listed in Figure 3.

In principle it is applicable for the homing phase and for final approach as well as for some last-minute abort manoeuvre.

The programme simulates the

- proportional navigation scheme, or constant bearing type of approach,
- pursuit course scheme, or variable bearing type of approach,
- combined schemes, with the above two in sequence.

But the programme also can simulate

- non-guided, free drifting fly-by and
- emergency manoeuvre to avoid collision.

The most important input data are shown in Figure 4. These obviously include

- spacecraft data
- orbital data
- directives concerning selected strategy and
- control parameters which finally determine the profile of the rendezvous process.

How they are involved in the simulation process is shown in Figure 5 in a simplified form. It deals with one control mode, namely with the braking manoeuvre of the chaser along its line-of-sight to the target.

This manoeuvre consists of the simultaneous reduction of range and range rate, to any specified value at a specified aiming point. Thus its simulation is comprised of the following:

- Orbital motion of the target with respect to an earth centered inertial reference frame on one hand, and orbit dynamics of the chaser's relative motion to the target using coordinates of a target centered inertial reference frame on the other hand.
- Transformation of relative Cartesian state variables into corresponding polar or spherical state variables.

- Generation of control variables. Whilst the previous two "boxes" on the right hand side are common to all modes, the control variables are specific to each mode. In the case shown only one control variable "AD" is necessary, representing the required axial deceleration to achieve final range rate " \dot{R}_E " at a range " R_E ".
- Finally a switching controller of an hysteresis type which steers the thrusting within specified tolerance values. These thrustings, in turn, are fed back into the chaser dynamics, i.e. into the corresponding differential equations.

4. CONTROL OF FINAL APPROACH

Coming now to some output of the subject simulation programme, it is obvious that within the limited time available, only very few results can be shown or discussed.

So let us pick out merely two cases out of the final approach phase:

- First - the mode of braking the chaser's approach towards the target, by axial thrusting along the line-of sight (LOS). This is the main part of the proportional navigation scheme, after rotation of LOS has been nearly brought to a stop.

When retaining this mode up to docking, a cooperative target spacecraft, in terms of orientation of its docking axis along the final LOS, is assumed.

- Secondly - the mode of rotating the LOS through a specified angle, e.g. at constant range from the target: this amounts to a chaser flight around the latter. This is necessary in case the direction of approach is not coincident with the docking axis of a fixed oriented target spacecraft.

For a very time-limited discussion of results, their illustration in a so-called "phase diagram" has proven itself most effective. It is a diagram showing range rate versus range.

Thus, in the first case discussed, the chaser's approach velocity is decreased as a function of its range from the target, Figure 6. This is done by using the comparison of necessary decelerations with achievable ones as criteria for switching the thrusters on and off. The necessary magnitudes represent boundaries of control channels and appear in the phase diagram - when defined as indicated - as parabolas.

It has been repeatedly confirmed that this method of controlling the braking impulses is convenient, flexible and efficient; care must be taken only to select the coefficients of the parabolas so as to

- be compatible with the state vector (R, \dot{R}) of the chaser at the initiation of the process,
- comply with the decelerations achievable by means of the thrust levels on board the chaser,
- result in the desired final values of R and \dot{R} .

Figure 7 shows a concrete case of approach to a target satellite flying on GTO at about a true anomaly of 150° , the chaser being on a nearby orbit starting its proportional navigation scheme at a distance of 1000 m off the target.

Having identified the excess in range rate at that range, the approach velocity is steeply reduced so as to enter the first control channel, compatible with the 400 N thruster performance.

The intermittent braking of range rate continues inside the channel until the second channel is achieved, which is compatible with the 40 N thrust level. The latter is necessary because the thrusting impulses get too short for the 400 N to deliver.

Here switching over to the lower level thrusters takes place on controller's command and the approach velocity is further reduced at a lower thrust level with increased pulse duration.

The dashed line, incidentally, represents free drift with no braking control.

It must be noted here that the convergence of the plot is no guarantee for clean docking conditions. If e.g. LOS stability is controlled at close proximity to the target based on its angular displacement rate, a chaser drifting around the target, with the risk of hitting it on the side, may occur. Therefore, a switch-over to LOS control based on lateral displacement, which we do at a range of about 200 m, is highly recommended.

As can be expected, the initial conditions of the chaser relative to the target are of importance.

Figure 8 shows for a circular 550 km target orbit the relatively slow approach of the chaser, it being about 1 km lower.

Doing nothing would cause the chaser to overtake the target at a range of 1000 m (dotted line) not achieving docking conditions.

Thus, an acceleration towards the target is necessary; the effects of this acceleration up to different range rate levels is shown in comparison.

Accelerating up to

- 1 m/s (long dashed line) seems inadequate, as the duration of approach becomes very long. The calculation was interrupted after about 1 hour flight equivalent time. So we don't know whether docking could be accomplished.
- 2 m/s (full line) results in docking conditions, smoothly achieved within the 40 N thrust control channel.
- 4 m/s (short dashed line) was too much to regain control of range rate, resulting in collision with an impact speed of 0,8 m/s.

This leads us to the possibility of simulating an abort of the final approach.

Figure 9 shows, on the left side, the case where the low thrust control channel, either intentionally or contingency-wise, has been suppressed. Leaving the high thrust channel, the chaser's controller allows free flight until a certain critical range, that is closest to the target, at which lateral thrust is still possible so as to result in a last minute circumventing of the target.

On the right side, both control channels are assumed to be out of function. Here again, the controller determines the last opportunity to avoid collision and therefore commands full braking and lateral thrust.

Finally, the case of rotating LOS, in terms of circum-flying the target, is shown in Figure 10. Here, range is kept constant during this manoeuvre. To this end, again, control channels are defined, representing, however, lateral deceleration limits within which the lateral thrustings take place.

Figure 11 shows at the left upper corner the effect of this rotation on the phase diagram plot. The rotation itself is, in fact, the single point on the $R = 0$ axis. The local vertical, there, indicates the re-acceleration towards the target, needed for completion of the approach after the rotation manoeuvre. It can be seen that – again – the level of this re-acceleration is not totally arbitrary, but rather must suit the available thrust level channel. Otherwise either a fly-by or a collision may occur.

5. A FEW CONCLUSIONS

Because of lack of time we end with only 3 conclusions:

- * The selected guidance schemes have been proven to be effective and realistic. Other optimum guidance schemes involve excessive complexity in on-board data processing, hardware and software implementation.
- * On high orbits the constant-bearing type approach, as well as the variable-bearing type, are rather unproblematic in achieving successful rendezvous.
- * The orbit curvature, with decreasing orbit altitude turns out to increasingly impede smooth rendezvous,
 - often requiring additional velocity stabilization control during proportional navigation,
 - rendering the pursuit course strategy inapplicable unless high thrust levels are available.

FIG. 1:

EUROPEAN SPACE AGENCY (ESA) SPONSORED RENDEZVOUS & DOCKING (RVD) STUDIES

MISSION	RVD-IN-LEO	RVD-IN-GTO	RVD-IN-GEO
OBJECTIVE	FLIGHT DEMONSTRATION OF RVD CAPACITY	ATTACHING APOGEE PROPULSION MODULE TO COMM-SAT PRIOR TO INJECTION INTO GEO	ASSEMBLY OF MODULARIZED GEO COMMUNICATION PLATFORM
INVOLVED LAUNCHER	SHUTTLE	ARIANE 3	ARIANE 4
TARGET ORBIT ALTITUDE	300 KM CIRCULAR 550 KM CIRCULAR	200/36000 KM	36000 KM CIRCULAR
RVD TARGET VEHICLE	"EURECA" RETRIEVABLE, FREE FLYING, PAYLOAD CARRYING PLATFORM	200 KG COMMUNICATION SATELLITE WITH NO APOGEE MOTOR	MODULARIZED GEO COMMUNICATION PLATFORM BEING BUILT UP
RVD CHASER VEHICLE	DEDICATED, SPECIALLY LAID OUT TEST VEHICLE	DEDICATED, ADAPTED APOGEE PROPULSION MODULE	COMMUNICATION PAYLOAD MODULE WITH ATTACHED PROPULSION KIT

FIG. 2:

THE RENDEZVOUS SIMULATION PROGRAMME STRUCTURE

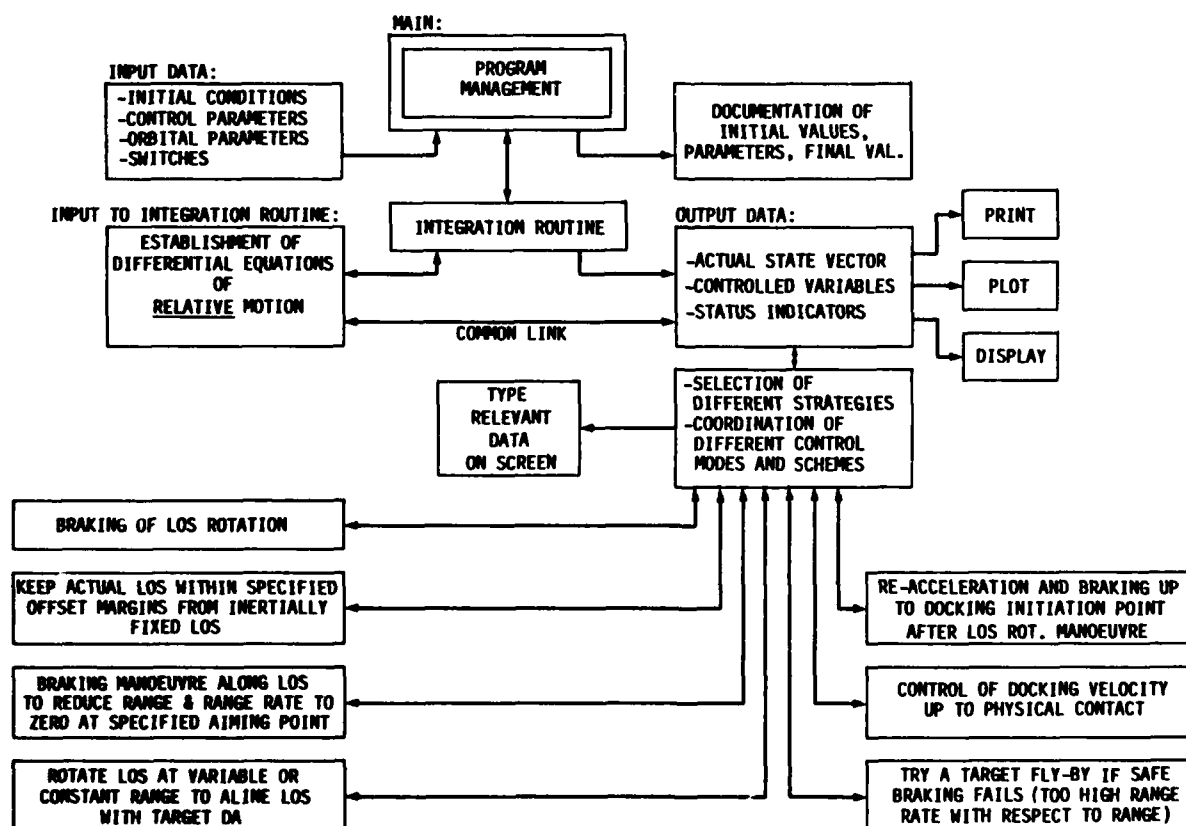


FIG. 3:

CAPABILITIES OF THE SIMULATION PROGRAMME

STRATEGY	CONSTANT BEARING FINAL APPROACH i.e. PROPORTIONAL NAVIGATION (PN)	<ul style="list-style-type: none"> o BRAKING OF LOS ROTATION o HOLDING LOS INERTIALLY FIXED o BRAKING RANGE RATE ALONG LOS o KEEPING CLOSING SPEED CONSTANT 	MODE
	VARIABLE BEARING FINAL APPROACH i.e. PURSUIT COURSE (PC)	<ul style="list-style-type: none"> o BRAKING RANGE RATE ALONG LOS o CONTROLLING ROTATION OF LOS o CONTROLLING RANGE RATE o HOLDING LOS INERTIALLY FIXED o KEEPING CLOSING SPEED CONSTANT 	
	COMBINED MODE OF FINAL APPROACH WITH PN AND PC IN SEQUENCE	COMBINATION OF ABOVE	
	CONSTANT BEARING HOMING	<ul style="list-style-type: none"> o RETAIN CONSTANT RANGE RATE o CONTROLLING LOS ROTATION 	
	EMERGENCY LAST MINUTE ABORT	<ul style="list-style-type: none"> o CHECK RANGE RATE VS. RANGE o PREDICT RELATIVE ORBITAL STATE AT ITP o ATTEMPT OF HOLD OR FLY-BY 	

FIG. 4 :

INPUTS AND CONTROL MAGNITUDES

SPACECRAFT DATA	
TARGET	- DOCKING AXIS ORIENTATION
CHASER	- INITIAL MASS - ORIENTATION - THRUSTER - THRUST LEVELS - SPECIFIC IMPULSES - DOCKING AXIS ORIENTATION

ORBITAL DATA	
TARGET ORBIT ELEMENTS	
INITIAL RELATIVE STATE VECTOR	

RENDEZVOUS STRATEGY	
SELECTED MODES/SEQUENCE	
IN-BETWEEN WAITING TIMES	
DOCKING VELOCITY	

RENDEZVOUS PROFILE	
CONTROL CHANNELS	- THRUST WEIGHING FACTORS - FINAL RANGE - FINAL RANGE RATE
LINE-OF-SIGHT (LOS)	- ROTATION ANGLE - ROTATION RATE - ROTATION THRESHOLDS - OFFSET BOUNDARIES
RANGE	- OF LOS ANGULAR RATE-TO-LATERAL OFFSET CONTROL SWITCH OVER - OF LOS ROTATION MANOEUVRE
RANGE RATE	- LOWER LIMIT OF RANGE RATE
DEVIATION FROM DOCKING AXIS	- LATERAL LIMITS - ANGULAR LIMITS

FIG. 5:

CHASER BRAKING ALONG LOS

(SIMULTANEOUS REDUCTION OF RANGE AND RANGE RATE TO ZERO OR SPECIFIC AIMING POINT)

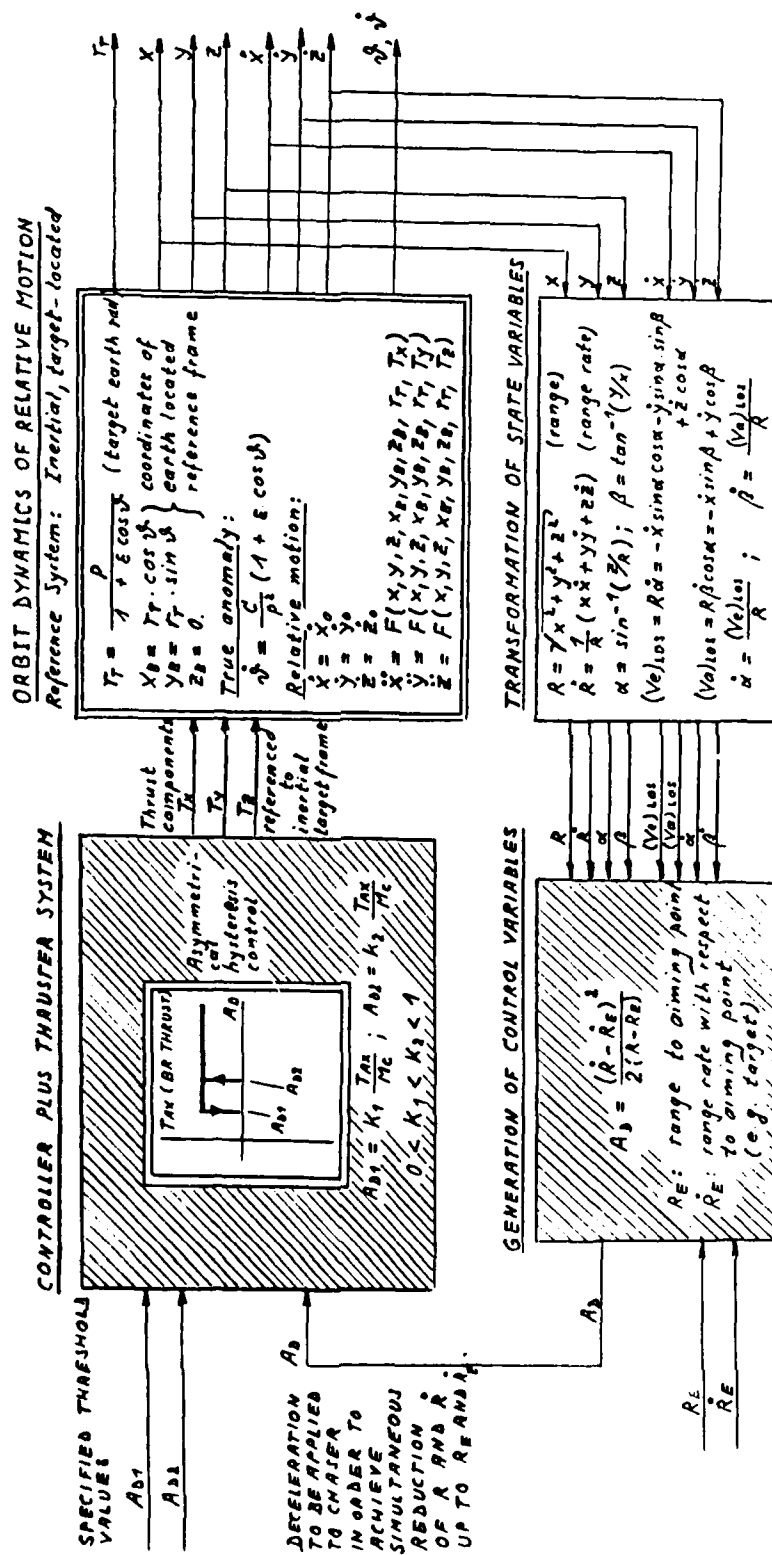


FIG. 6:

CHASER BRAKING ALONG LOS

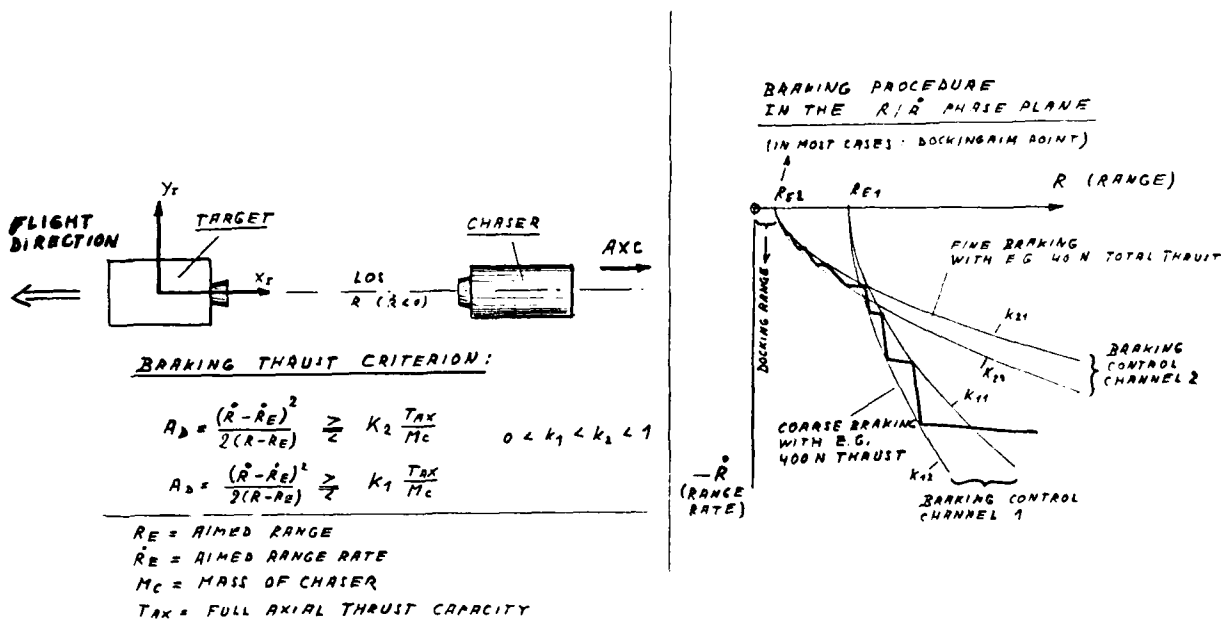


FIG. 7:

TWO-LEVEL-THRUST APPROACH IN THE PHASE DIAGRAMME

TARGET ORBIT: 200/35800 KM

CHASER INITIAL RANGE: 1000 M

CHASER MASS: 2000 KG

THRUST LEVEL, AXIAL : 400 N

LATERAL: 40 N

FINAL RANGE OF

400 N CONTROL CHANNEL: 200 M

40 N CONTROL CHANNEL: 0 M

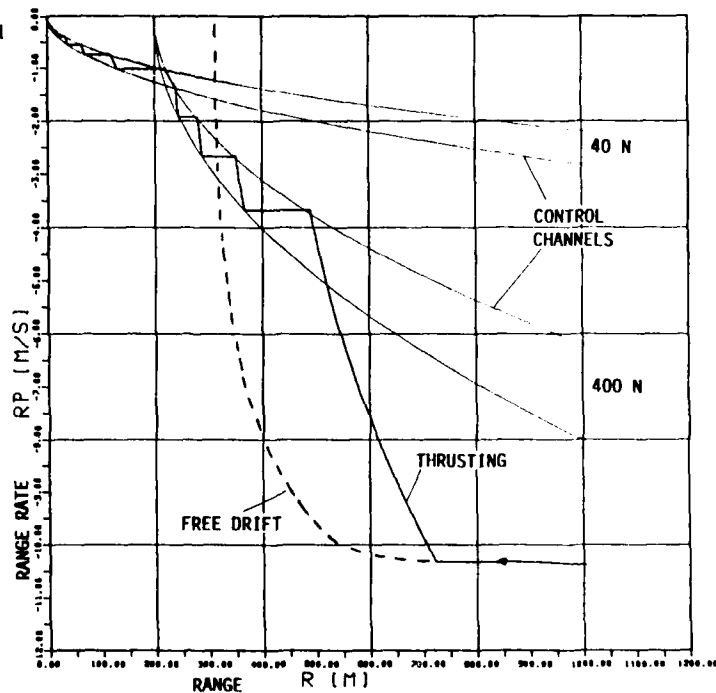


FIG. 8:

EFFECTS OF ACCELERATION LEVEL

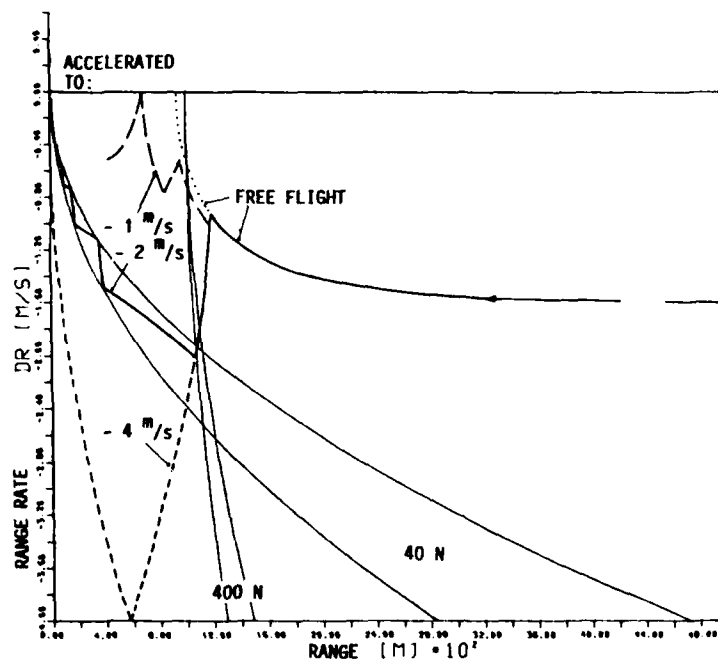


FIG. 9:

FINAL APPROACH ABORT

TARGET ON 200/35800 KM ORBIT

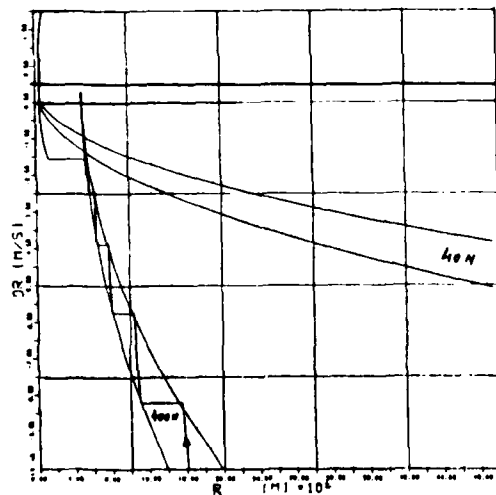
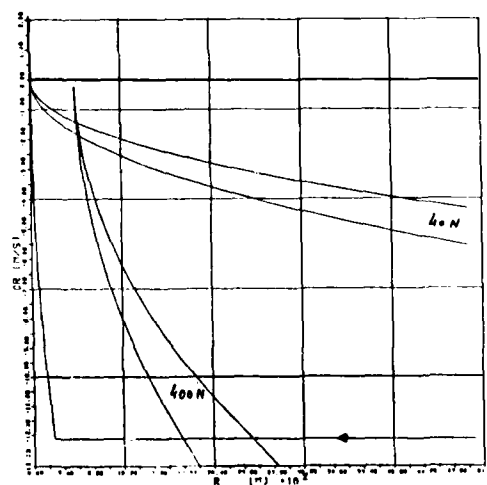
SUPPRESSION OF 40 N THRUST CHANNEL
FLY-BY AT 100 MSUPPRESSION OF BOTH CONTROL CHANNELS
FLY-BY AT 25 M

FIG. 10:

ROTATE LOS TO DESIRED ORIENTATION AT CONSTANT RANGE

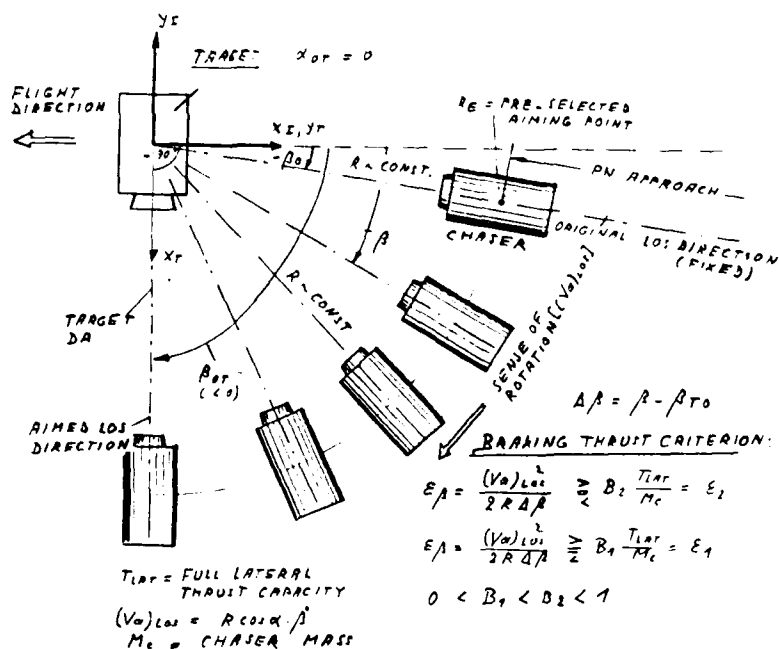
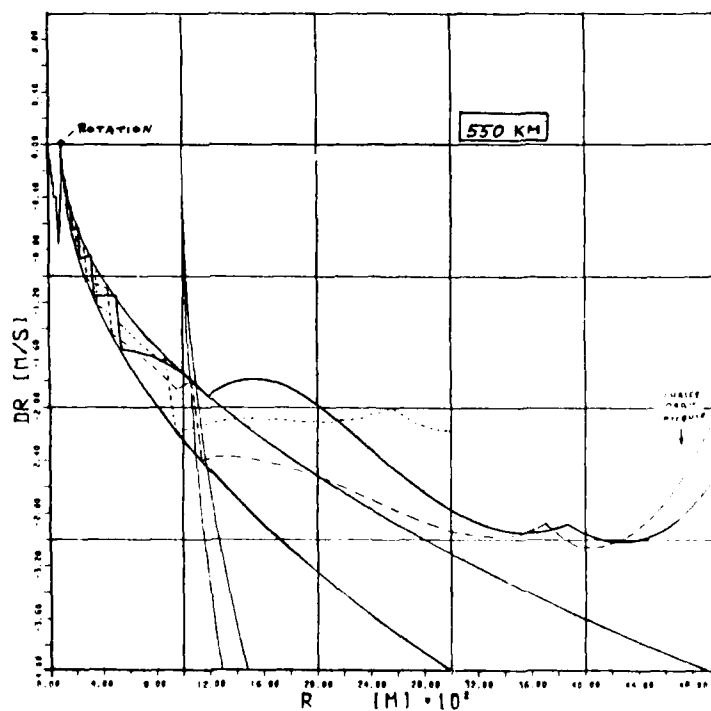


FIG. 11:

FINAL APPROACH WITH LOS ROTATION



OPTIMISATION DES MANOEUVRES D'UN SATELLITE D'OBSERVATION EN VUE D'UN SURVOL RAPIDE D'UN POINT DU GLOBE

par

A.M.Mainguy, C.Aumasson et S.Laffy
ONERA
B.P.72
92322 Chatillon Cedex
France

Un satellite d'observation de la Terre est placé de façon nominale sur une orbite de surveillance quasi polaire (héliosynchrone) circulaire et d'altitude telle que l'entrelacement des traces au sol soit suffisamment régulier et serré. Ceci assure une couverture convenable du globe terrestre sur une période de quelques jours.

Pour des raisons de haute surveillance, il peut s'avérer nécessaire de modifier cette orbite et de placer la satellite sur une orbite circulaire "géosynchrone" permettant d'observer un point précis du globe de façon répétitive une fois par jour et pendant un temps donné. Cette manoeuvre doit être effectuée le plus rapidement possible, en moins de deux jours, par exemple. Les deux orbites étant d'altitudes voisines, il est possible de s'affranchir du maintien de la contrainte d'héliosynchronisme et de résoudre le problème en supposant les deux orbites de même inclinaison.

La difficulté du problème réside dans le fait que la manoeuvre doit être effectuée sur une courte durée de temps avec des réservoirs de capacité limitée et des propulseurs à faible niveau de poussée. Il faut procéder par étapes pour pouvoir résoudre le problème dans sa généralité.

Une première étude est faite avec des hypothèses simplificatrices portant à la fois sur les contraintes physiques (poussée impulsionnelle) et sur les équations mathématiques (linéarisation des équations autour de l'orbite nominale). Il est ainsi possible de construire analytiquement les solutions optimales ou sous optimales (à poussée tangente) et d'évaluer les ordres de grandeur des coûts des manoeuvres, en fonction de la position du point survolé et du temps disponible pour réaliser l'objectif. Dans le cas de durée assez longue, le coût est légèrement supérieur à celui d'un simple transfert sur orbite géosynchrone. Mais pour des durées plus courtes, le coût du rendez-vous devient considérable, les accroissements de vitesse à fournir ne peuvent plus être délivrés dans une durée compatible avec l'approximation impulsionnelle et dans certains cas conduisent à une consommation dépassant la capacité maximale des réservoirs. Malgré la relative proximité des orbites initiale et finale, les orbites intermédiaires sont significativement différentes. En conséquence, les hypothèses simplificatrices ne sont plus valables.

Une deuxième étude s'appuie sur des hypothèses plus réalistes, en prenant en compte le faible niveau de poussée, et donc l'étalement des arcs propulsés. En supposant encore les équations d'évolution linéarisées et la poussée tangente, le problème se présente comme un problème d'optimisation paramétrique avec contraintes, dont les inconnues sont les instants d'allumage et d'extinction des moteurs. Il ne peut être résolu qu'en utilisant un algorithme numérique itératif mais il apparaît alors une difficulté liée à l'initialisation de la méthode. En effet, une simple transcription des résultats de la première étude ne permet pas toujours d'aboutir à une convergence satisfaisante de l'algorithme; une méthode originale d'initialisation récursive a été développée. Les résultats obtenus permettent de mettre en évidence la dégradation du coût en fonction de l'étalement, d'autant plus importante que le coût lui-même est élevé. Indépendamment de la capacité maximale des réservoirs, certains points sont même impossibles à survoler dans un bref délai, car l'étalement des arcs de poussée est trop grand, et impossible à répartir sur le nombre de tours disponible pour le survol des points (les points limites du domaine accessible sont associés à une poussée ininterrompue). Une première approche des manoeuvres réalisables a ainsi été obtenue, mais ce n'est qu'une approche puisque les équations du mouvement ont été simplifiées.

Une troisième étude permet d'aborder le problème réel sans simplification. Il s'agit d'une optimisation fonctionnelle où les commandes à déterminer sont le module de la poussée et son orientation dans le repère orbital local. La méthode numérique itérative utilisée est plus complexe que la précédente quoique basée sur le même principe. Il a été vérifié que la méthode converge même lorsque l'initialisation est effectuée avec une poussée identiquement nulle et vers une solution de type tout ou rien, compatible avec le principe du maximum de Pontryagin. Ceci a permis de vérifier le bon ordre de grandeur obtenu dans le cadre de l'étude précédente et de réajuster les instants d'allumage et d'extinction des moteurs. La possibilité d'optimiser également l'orientation de la poussée montre que la poussée tangentielle est quasi optimale sauf pour des survols effectués en moins de 2 tours où la poussée librement orientable peut s'avérer intéressante.

AN APPROACH TO ATTITUDE CONTROL MANEUVERS WHICH HELPS MAINTAIN ORBIT INCLINATION

Dr. W. Michael Bowles
Senior Scientist

and

Mark R. Altobelli
Technical Staff

Space and Communications Group
Hughes Aircraft Company
El Segundo, CA

HUGHES PROPRIETARY

This document contains proprietary information and except with written permission of Hughes Aircraft Company, such information shall not be published, or disclosed to others, or used for any purpose, and the document shall not be duplicated in whole or in part.

I. INTRODUCTION

The procedure described here is useful for a spin stabilized satellite spinning around an axis normal to its orbit plane and using an offset axial thruster for attitude control. For such a satellite, the procedure arranges attitude control maneuvers so that such maneuvers help maintain orbit inclination.

The important features of the satellite are that it spins with its spin axis normal to the orbit plane and it uses offset axial thrusters to control attitude and to impart velocity along the spin axis. This configuration is sketched in Figure 1. Both axial thrusters are fired simultaneously to impart a velocity along the spin axis. To alter the

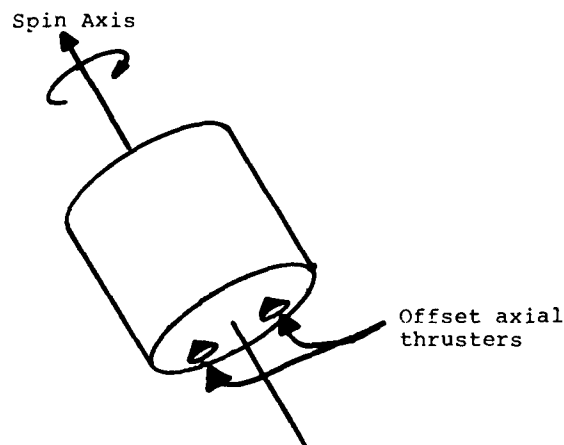


Figure 1. Spinning Satellite Using Offset Axial Thrusters

spin axis attitude one thruster is pulsed. If the pulse is short enough that the satellite does not rotate through a very large angle, then the effect is to apply an external torque to the satellite and thus to precess the spin axis. By repeatedly pulsing the thruster the spin axis can be precessed by any desired amount. This manner of achieving a spin axis reorientation also results in a velocity change in the spin axis direction. The object is to make sure that whenever it is necessary to reorient the spin axis, the reorientation can be done so that the attendant velocity change is also helpful.

A satellite in a circular orbit will not remain in that orbit indefinitely if left alone. Among other changes that the orbit undergoes, it tilts relative to the starting orbit. In order to stay close to a desired orbit the accumulated tilt must be corrected out periodically. Figure 2 helps indicate how the orbit tilt correction may be carried out. When the satellite in its tilted orbit crosses the desired orbit going down, a velocity change is introduced normal to the orbit. This swings the velocity vector for the satellite from that connected with the tilted orbit to that appropriate for the desired orbit.

Besides not remaining in a fixed circular orbit a satellite will not hold a fixed attitude. The attitude drifts principally to the influence of solar torque. For a spinning satellite, solar torque results in a precession of the satellite around the line between the satellite and the sun. Periodically, the attitude must be corrected.

On the spinning satellite with offset axial thrusters, altering attitude also imparts some velocity normal to the orbit. If all the attitude trim maneuvers are done at the descending node, then the axial velocity tends to help maintain orbit inclination.

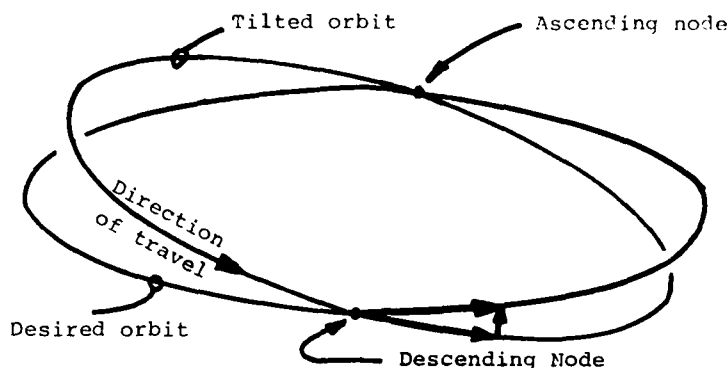


Figure 2. Geometry of Orbit Tilt Correction

In some instances it is not possible to do all of the attitude control maneuvers at the descending node. In particular, when the precession rate gets so large that the precession accumulated in one day exceeds the bounds of acceptable pointing for the satellite spin axis, then it is not possible to wait for one day between maneuvers. How to arrange the attitude control maneuvers in this case is the subject of this paper.

II. DESCRIPTION OF SOLUTIONS

The problem is to determine an optimum strategy for performing attitude control maneuvers when the combination of pointing tolerance and precession rate force corrections to be made more often than once per day. An optimum strategy is one which maintains spin axis attitude within acceptable bounds and which arranges attitude control maneuvers in such a way as to achieve the greatest beneficial effect on orbit inclination. The optimum strategy can be found through intuitive argument and more rigorous proof of optimality follows the intuitive argument quite closely.

These are two state variables of interest here. One is the angle of rotation of the satellite around the line from the satellite to the sun. This angle will be denoted by θ_r . Orbit normal will be $\theta_r = 0$. The other is the angle of the satellite in its orbit. Let θ_0 denote the angle between a line from the earth's center to the ascending node of the orbit and a line from the earth's center to the satellite.

The sequence of plots shown in Figure 3 will help explain the strategy for optimizing attitude control maneuvers. Suppose first that there is no restriction on the excursion for the spin axis. Figure 3a shows how the attitude θ_r would behave as a function of the satellite's position in orbit θ_0 if all the corrections were done at the descending node. In this case the attitude starts from zero, suppose, and increases until the descending node is reached. At that point a correction equal to the daily precession rate is introduced. Until the next passage through the descending node, the attitude is allowed to drift. Suppose now that the attitude cannot be allowed to drift for an entire day. Suppose that the precession in one orbit is B . Suppose also that to ensure proper performance of the satellite, the rotation of the spin axis around the sun line must be constrained to within $\pm \delta$ of orbit normal. Figure 3b shows a strategy for performing attitude control maneuvers which keeps $|\theta_r| \leq \delta$ for the case where $\delta = B/4$. In Figure 3b the attitude begins at zero. It cannot be allowed to drift until the descending node as was the case in Figure 3a. The upper bound on the spin axis attitude deviation is reached at the antinode $\theta_0 = \pi/2$. Some correction must be done at $\theta_0 = \pi/2$. The decision is made to introduce a correction of $B/4$. When this is done the attitude θ_r will just drift to the upper bound δ at the descending node ($\theta_0 = \pi$). This is a good choice because it allows for a correction of maximum size to be done at the descending node and thus benefits orbit tilt. The total correction done during the orbit must equal the total precession during the orbit. It makes sense to do as much of this correction as possible at the descending node. The corrections done at the antinode cancel one another's contribution to orbit tilt. The efficiency of this strategy is then 50% in that 50% of the linear velocity imparted by the attitude control maneuver is helpful for orbit tilt correction.

It is possible for firings done away from the descending node to benefit orbit tilt. Suppose that a linear velocity of ΔV is imparted at orbit angle θ_0 . This ΔV certainly tilts the orbit by some amount.

If a velocity change ΔV imparted in the spin axis direction at the descending node tilts the orbit plane in the desired direction by α degrees, then ΔV imparted at orbit angle θ_0 degrees tilts the orbit plane in the desired direction by $-\delta \cos \theta_0$ degrees. Attitude maneuvers done at the antinodes ($\theta_0 = \pi/2$ or $\theta_0 = 3\pi/2$) tilt the orbit in a direction orthogonal to the direction in which correction is required. Corrections done at the ascending node make the tilt worse. It is best if the attitude maneuvers are gathered close to the descending node.

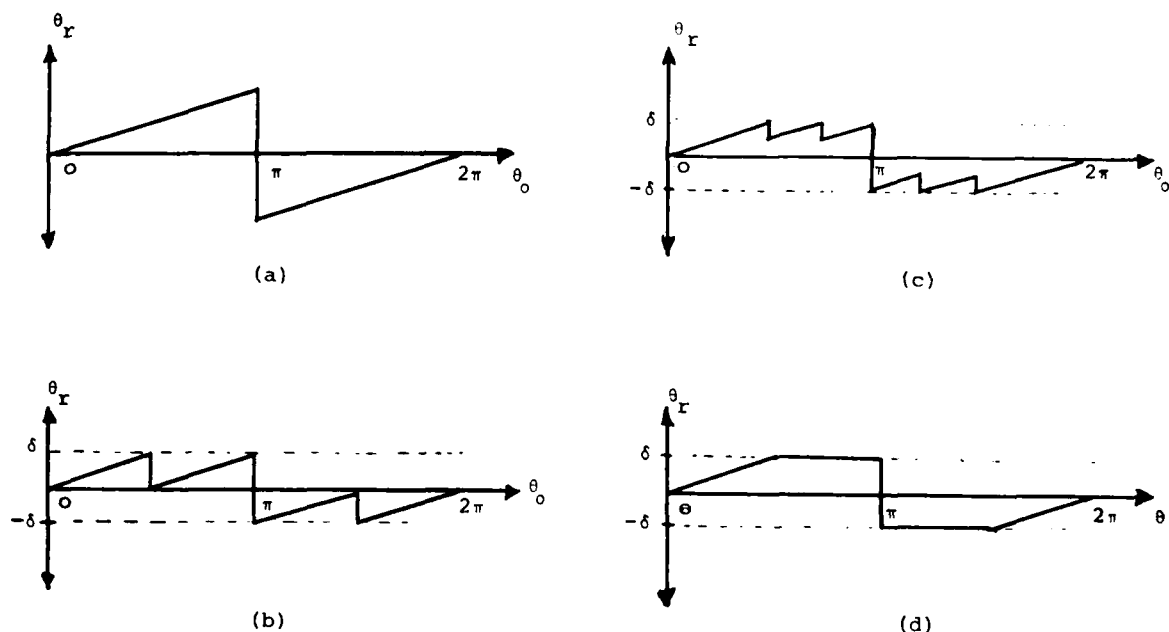


Figure 3. Geometry of Attitude Maneuvers

The strategy depicted in Figure 3b can be improved upon by using this guideline. Instead of doing a maneuver of $B/4$ at the antinode where the threshold δ is first reached suppose that a maneuver of $B/8$ were done. The threshold would be reached again at $\theta_o = 3\pi/4$ and another maneuver of $B/8$ could be done. A strategy following this outline is depicted in Figure 3c. The two maneuvers of $B/8$ which have been moved to $\theta_o = \pi \pm \pi/4$ benefit orbit tilt and so the maneuver of Figure 3c is an improvement over the maneuver in Figure 3b.

The maneuvers could be further subdivided and in the limit the result would look like the graph in Figure 3d. In the limiting case the strategy is to do nothing until the threshold δ is hit. Thereupon, a series of small attitude corrections is done. These small attitude corrections just offset the precession and the attitude offset remains constant at δ until the descending node is reached. At the descending node a correction of 2δ is introduced. Past the descending node, another series of small corrections is done. These corrections hold the attitude at $\theta_r = -\delta$. The attitude is at $\theta_r = \delta$ over the same angle prior to the descending node as it is at $\theta_r = -\delta$ after the descending node. This fact and the value of the threshold δ and the precession rate B determine at what angle θ_o the firing ceases and the attitude begins to drift.

An efficiency for this strategy can be defined to be the ratio of the benefit derived by using the strategy to the benefit achieved by doing the entire attitude control maneuver at the descending node. To compute this efficiency, first recognize that the orbit angle α through which the satellite travels while the attitude precesses from $-\delta$ to $+\delta$ satisfies

$$\frac{B\alpha}{2\pi} = 2\delta \quad (1)$$

In this equation B is the precession rate in radians of precession per orbit, α is an orbit angle in radians and δ is the allowable deviation of the attitude from orbit normal. According to the strategy the satellite exerts a low-level thrust which just offsets the precession rate $B/2\pi$ when its orbit angle is between $\frac{\alpha}{2}$ and $2\pi - \alpha/2$ or when θ_o is between $2\pi \frac{\delta}{B}$ and $2\pi - 2\pi \frac{\delta}{B}$. The benefit derived by this strategy is

$$\int_{2\pi \frac{\delta}{B}}^{2\pi(1 - \frac{\delta}{B})} -\frac{B}{2\pi} \cos \theta_o d\theta_o + 2\delta \quad (2)$$

The integral term arises from the low-level thrusting and the 2δ term is from the impulsive attitude change at the descending node. If the precession for an entire orbit were taken out at the descending node, then the benefit would be B . The efficiency of the strategy outlined can be defined as the ratio of these two quantities. Denote the efficiency by E then

$$E = \frac{- \int_{\frac{2\pi\delta}{B}}^{2\pi(1-\frac{\delta}{B})} \frac{B}{2\pi} \cos\theta_0 d\theta_0 + 2\delta}{B} \quad (3)$$

$$E = 2 \frac{\delta}{B} + \frac{\sin \frac{2\pi\delta}{B}}{2\pi} \quad (4)$$

This equation is only valid for $B \geq 2\delta$. For $B < 2\delta$ the efficiency is 1.0. Notice that this expression only depends on the ratio δ/B . Figure 4 is a plot of E versus δ/B . As a single point comparison consider the example where $B = 4\delta$. The strategy shown in Figure 2b gave an efficiency of 0.5. By comparison the optimal strategy yields an efficiency of 0.82.

It is fairly straight forward to show that the strategy proposed satisfies first order necessary conditions. That is, that small changes in the timing or magnitude of the attitude control maneuvers decrease the efficiency of the maneuver. Figure 5 illustrates the optimal strategy and some small changes point to it. These changes will be shown to decrease the efficiency. The optimal strategy is, as has been described. The attitude starts at zero at the ascending node ($\theta_0 = 0$). The attitude is allowed to drift until the upper threshold ($\theta_r = \delta$) is reached ($\theta_0 = 2\pi\delta/B$). After that point the attitude is maintained at the upper threshold until the descending node is reached ($\theta_0 = \pi$). At the descending node the attitude is changed impulsively from $\theta_r = \delta$ to $\theta_r = -\delta$. The attitude is then maintained at the lower threshold until that point in the orbit from where the orbit will just drift back to zero at the ascending node.

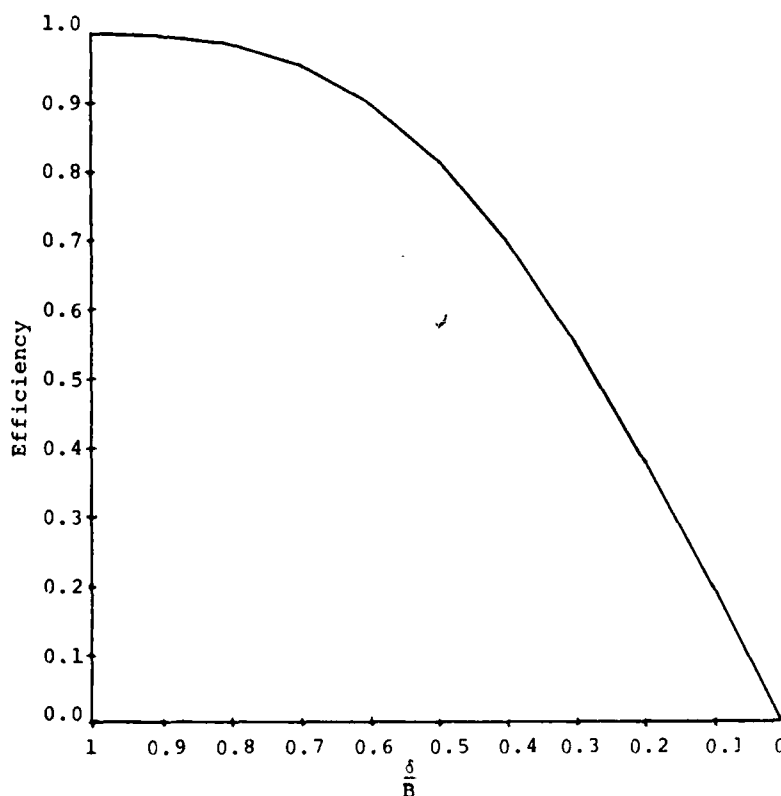


Figure 4. Efficiency of Optimal Maneuver Strategy

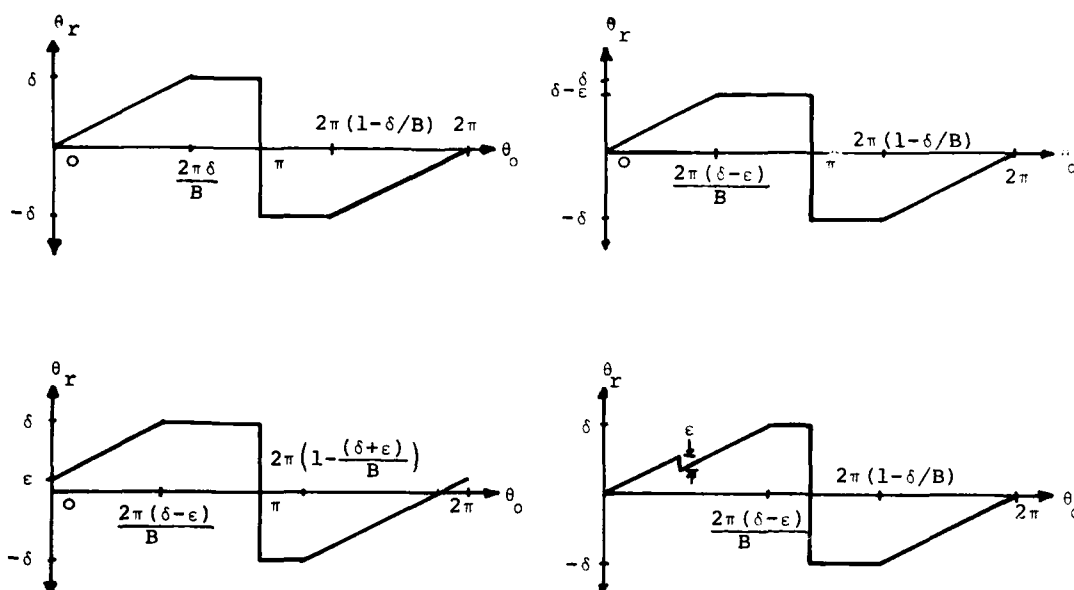


Figure 5. Optimal Strategy and Perturbations

First, as illustrated in Figure 5b, suppose that the attitude does not start at $\theta_r = 0$ but at $\theta_r = \epsilon$ where ϵ is a small number. Denote the efficiency for this perturbed case by E' then the change in efficiency

$$E' - E = \frac{-\int_{\frac{2\pi(\delta-\epsilon)}{B}}^{2\pi\left[1 - \frac{(\delta+\epsilon)}{B}\right]} \frac{B}{2\pi} \cos\theta_o d\theta_o + 2\delta}{B} - \frac{-\int_{\frac{2\pi\delta}{B}}^{2\pi\left(1 - \frac{\delta}{B}\right)} \frac{B}{2\pi} \cos\theta_o d\theta_o + 2\delta}{B} \quad (5)$$

$$E' - E = \frac{-\int_{\frac{2\pi(\delta-\epsilon)}{B}}^{\frac{2\pi\delta}{B}} \frac{B}{2\pi} \cos\theta_o d\theta_o + \int_{2\pi\left[1 - \frac{(\delta+\epsilon)}{B}\right]}^{2\pi\left(1 - \frac{\delta}{B}\right)} \frac{B}{2\pi} \cos\theta_o d\theta_o}{B} \quad (6)$$

Consider the behavior of $E' - E$ as a function of ϵ . Denote this dependence by setting

$$g(\epsilon) = E' - E \quad (7)$$

Then straight forward computation yields

$$\begin{aligned} g(0) &= 0 \\ g'(0) &= 0 \\ g''(0) &= -\frac{4\pi}{B^2} \sin\left(2\pi \frac{\delta}{B}\right) \end{aligned} \quad (8)$$

Since $\delta/B \leq \frac{1}{2}$ then $g''(0) \leq 0$ and $g''(0) < 0$ for $\delta/B < \frac{1}{2}$. Therefore, the effect of making the starting attitude slightly different from zero is to decrease the efficiency of the maneuver.

Next suppose, as shown in Figure 5c, that the attitude is not allowed to reach the threshold δ but is stopped at $\delta - \epsilon$. In this case

$$E' - E = \frac{- \int_{\frac{2\pi(\delta - \epsilon)}{B}}^{\frac{2\pi\delta}{B}} \frac{B}{2\pi} \cos\theta_O d\theta_O - \epsilon}{B} \quad (9)$$

Again let $g(\epsilon) = E' - E$ then $g(0) = 0$ and

$$g'(0) = - \frac{1 + \cos \frac{2\pi\delta}{B}}{B} \quad (10)$$

In the last case $g'(0)$ was zero and $g''(0)$ was negative. This was a reflection of the fact that in that case the initial attitude could be increased or decreased. Mathematically speaking, in that case ϵ could be either positive or negative. In this case, by contrast, ϵ can only be positive. If ϵ were negative, then the constraint that the attitude stay between $+\delta$ and $-\delta$ would be violated.

The last perturbation to the optimal strategy assumes that the attitude is changed by ϵ at orbit angle θ where $0 \leq \theta \leq \frac{2\pi\delta}{B}$. Assume first that $\epsilon > 0$. In this case the change in efficiency is given by

$$E' - E = \frac{\int_{\frac{2\pi(\delta + \epsilon)}{B}}^{\frac{2\pi\delta}{B}} \frac{B}{2\pi} \cos\theta_O d\theta_O - \epsilon \cos\theta}{B} \quad (11)$$

As before let $g(\epsilon) = E' - E$ then $g(0) = 0$ and

$$g'(0) = \frac{1}{B} \left(\cos \frac{2\pi\delta}{B} - \cos\theta \right) \quad (12)$$

Since $\pi \geq \frac{2\pi\delta}{B} \geq \theta \geq 0$ then $g'(0) \leq 0$.

On the other hand if $\epsilon < 0$ then

$$E' - E = \frac{- \int_{\frac{2\pi(\delta - \epsilon)}{B}}^{\frac{2\pi(1 - \frac{\delta}{B})}{B}} \frac{B}{2\pi} \cos\theta_O d\theta_O + 2\delta + \epsilon \cos\theta}{B + 2\epsilon} \quad (13)$$

$$- \frac{\int_{\frac{2\pi\delta}{B}}^{\frac{2\pi(1 - \frac{\delta}{B})}{B}} \frac{B}{2\pi} \cos\theta_O d\theta_O + 2\delta}{B}$$

Clearly $E' - E = 0$ when $\epsilon = 0$

To first order in ϵ

$$E' - E = \frac{\epsilon \cos \frac{2\pi\delta}{B} + \epsilon \cos\theta}{B} + E \left(1 - \frac{2\epsilon}{B}\right) - E \quad (14)$$

$$= \frac{\epsilon \cos \frac{2\pi\delta}{B} + \epsilon \cos\theta}{B} - \frac{2\epsilon}{B} E \quad (15)$$

This will be negative if

$$- \cos \frac{2\pi\delta}{B} - \cos\theta \leq 2E \quad (16)$$

But

$$- \cos \frac{2\pi\delta}{B} - \cos\theta < -2 \cos \frac{2\pi\delta}{B}$$

and

$$E > \frac{2\delta - (B-2\delta) \cos \frac{2\pi\delta}{B}}{B} \quad (17)$$

$$= 2 \frac{\delta}{B} - \left(1 - 2 \frac{\delta}{B}\right) \cos \frac{2\pi\delta}{B} \quad (18)$$

and

$$- \cos \frac{2\pi\delta}{B} \leq \frac{2\delta}{B} - \left(1 - \frac{2\delta}{B}\right) \cos \frac{2\pi\delta}{B} \quad (19)$$

or

$$0 \leq 1 + \cos \frac{2\pi\delta}{B} \quad (20)$$

Since this inequality is obviously true, the ones preceding it are true also. Therefore, the perturbation shown in Figure 5d decreases the efficiency of the maneuver.

The maneuver described as optimal has been shown to satisfy first order necessary conditions and some second order necessary conditions. Furthermore, it satisfies intuition. There is a practical difficulty with it however.

The torque causing the satellite to precess is usually so small that the thrusters cannot fire at a low enough level to just offset that torque. The attitude maneuvers tend to be a series of impulsive maneuvers of some minimum size instead of a continuous low level firing. An approach which acknowledges this practical reality, but which is close to optimal is discussed in the next section.

III. ALGORITHM FOR QUANTIZED ATTITUDE MANEUVERS

Basically, the attitude control strategy is a sequence of four types of maneuvers. The first type is no thrusting. The attitude drifts without correction in this section. The second is thrusting to just keep the attitude on the upper threshold. The third is an impulsive maneuver at the descending node. The fourth is thrusting to just keep the attitude on the lower threshold. After the fourth section, the strategy goes back to the first section and does not exert any correction. Thereafter the sequence is repeated.

Now suppose that the only available attitude maneuver is an impulsive attitude change of some minimum amplitude σ . Suppose also that the satellite begins with zero attitude at the ascending node. If the optimal strategy is followed then the attitude precesses until the upper threshold is reached. At that point some attitude maneuver must be executed. The minimum is an impulse of magnitude σ . The attitude is then below the threshold by σ and can be allowed to drift until the threshold is again reached. Another impulsive change of σ is made. This sequence continues until the descending node is

reached. At the descending node the desire is to execute a maneuver which will carry the attitude down to the lower threshold δ . Since the only maneuver which can be executed is an impulsive change of magnitude σ then a rapid succession of changes are made. When the attitude passes $-(\delta-\sigma)$ firing must stop. Another pulse would carry the attitude outside the band $\pm \delta$. The attitude is allowed to precess until the threshold $-(\delta-\sigma)$ is reached.

At this point a correction of σ can be executed. This correction will just take the attitude down to the lower limit. The attitude is then allowed to precess again until the level $-(\delta-\sigma)$ is reached. The sequence continues until the time where correction ceases and the attitude is allowed to drift. The attitude then drifts up to the upper threshold δ and the operation begins afresh.

What has developed is a procedure whereby the attitude is continuously compared to a switching threshold. As long as the attitude is lower than the threshold, it is allowed to drift. If the attitude is greater than or equal to the threshold, a maneuver is executed. The threshold value is switched between $+\delta$ and $-(\delta-\sigma)$. The value $-(\delta-\sigma)$ is used for the threshold from the descending node until the point where the optimal strategy would allow the attitude to drift. During the remaining portion of the orbit the value δ is used. Figure 6 shows how the attitude might vary using this algorithm.

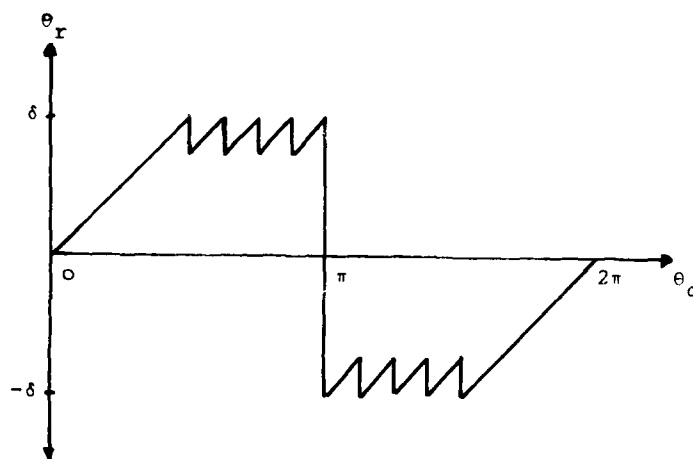


Figure 6. Attitude History Using Impulsive Control

Extensive numerical computations have been done to estimate the fuel savings that this algorithm will yield for the Intelsat VI satellite being built by an international team headed by Hughes Aircraft Company for the International Telecommunications Satellite Organization (INTELSAT). This satellite is designed to provide major international communications traffic links. The satellite will be positioned in geosynchronous equatorial orbit. It is required that the attitude drift be kept within $\pm 0.04^\circ$. Precession of the satellite is principally due to solar torque. The solar torque precession rates vary over the ten-year design life of the satellite. This variation is principally due to changing spacecraft mass as a result of propellant expenditure. In addition, the precession rate varies over the year because the equatorial orbit changes its aspect angle to the sun.

The precession rate is fairly constant over a day, but there is some variation in the rate due to the changing aspect angle of the satellite's earth-pointing antennas. The maximum precession rate is about 0.14° per day. The minimum impulsive attitude change about 0.02° . Numerical calculations indicate that, in all, about 14.2 kg of propellant will be required for attitude maneuvers on Intelsat VI during its ten-year life. Of this 14.2 kg, 13.0 kg contributes directly to control orbit inclination.

IV. SUMMARY

An algorithm for attitude control has been presented that derives a benefit for orbit inclination. This algorithm is useful for a spinning satellite using an offset axial thruster to perform attitude maneuvers. Depending on the circumstances of the mission, significant fuel savings may be derived by the use of this algorithm.

CHARACTERIZATION OF UNCERTAINTY FOR LARGE SPACE STRUCTURE CONTROL PROBLEMS

John C. Doyle and Joseph E. Wall

Honeywell, Inc.
Systems and Research Center
2600 Ridgway Parkway
Minneapolis

SUMMARY

This paper reviews a new formulation of the feedback control problem that captures both its performance and robustness aspects. The basic analysis tool in this formulation is the structured singular value. The methods are potentially applicable to Large Space Structure Control Problems since they allow for uncertainty in a very natural way.

1. INTRODUCTION

The basic requirement of feedback systems is to achieve certain desired levels of performance and also to be tolerant of uncertainties. Performance levels concern such things as command following, disturbance rejection, sensitivity, etc., while uncertainty tolerances deal with the inevitable differences which exist between a physical plant and its mathematical model. As discussed in various textbooks and references, these two aspects of the feedback problem lead to fundamental tradeoffs and compromises which motivate the entire body of feedback theory. ([1], [2])

An essential difficulty in the theory has been to capture both the performance and uncertainty aspects of feedback in a single problem statement. A recent problem formulation has been proposed which captures both aspects of feedback under the umbrella of what is called the "block-diagonal bounded perturbation problem." The solution to this problem involves a generalization of the ordinary singular value decomposition. It provides a reliable, nonconservative measure to determine whether both the performance and robustness requirements of a feedback loop are satisfied. This measure is called the structured singular value (SSV) and serves as the essential analysis tool. One of the major goals of this paper is to disseminate this problem formulation. ([4], [5])

The second goal of this paper is to briefly indicate how the structured singular value methods apply to Large Space Structures (LSS) Control Problems. LSS problems have performance requirements for pointing, tracking, and shape control. These performance requirements must be maintained in the face of significant uncertainty in the form of disturbances, noises, and uncertainty in the structure itself and its actuators and sensors. Maintenance of performance in the face of uncertainty will be referred to as robust performance.

The SSV methods outlined in this paper provide the first reliable technique for analyzing the performance of a system in the presence of both uncertain inputs (disturbances and noises) and uncertain plant dynamics. Previous methods, based on covariance or singular value analysis, evaluated either nominal performance or robustness of stability but not robust performance. Thus the SSV methods provide a promising alternative for LSS control problems.

2. MATHEMATICAL PRELIMINARIES

2.1 Topics from Functional Analysis

The foundation for the development in this paper is provided by a review of a number of basic concepts from functional analysis. The first of these is the normed linear space L_2^{rxm} . This function space is the collection of all rxm -dimensional functions which are square integrable on \mathbb{R} . An inner product for any two functions x and y in L_2^{rxm} is defined as

$$\langle x, y \rangle = \int_{-\infty}^{+\infty} \text{Tr}[y^H(t)x(t)] dt \quad (2.1)$$

The norm associated with this inner product is

$$\|x\|_2 = \langle x, x \rangle^{1/2} \quad (2.2)$$

Elements of L_2^{rxm} have finite norm and are equivalence classes of functions differing on sets of measure zero.

AD-AD P003 394

The chief limitation of the space L_2^{rxm} for control system analysis is that it contains no unstable functions, i.e., functions with $\|x\|_2 = \infty$. This can be remedied by introducing the extended normed linear space L_{2e}^{rxm} . This space is the collection of all functions which are square integrable on all finite intervals of R . More precisely, we introduce the truncated norm as

$$\|x\|_{2,\tau} = \left\{ \int_{-\tau}^{\tau} \text{Tr}[x^H(t)x(t)] dt \right\}^{1/2} \quad (2.3)$$

Then L_{2e}^{rxm} contains all functions $x: R \rightarrow C^{rxm}$ which satisfy $\|x\|_{2,\tau} < \infty$ for all τ . Functions such as $x(t) = e^t$ are included in L_{2e}^{rxm} , for example, while functions such as $x(t) = \tan(t)$ are not. Note that all elements of L_2^{rxm} are included in the extension L_{2e}^{rxm} and have the property that $\|x\|_{2,\tau} \rightarrow \|x\|_2$ as $\tau \rightarrow \infty$.

The other function space we will need is L_{∞}^{rxm} . This space consists of all functions which are measurable and essentially bounded. It is a normed linear space with norm

$$\|x\|_{\infty} = \text{ess sup } \sigma(x(t)) \quad (2.4)$$

No distinction will be made between functions differing over sets of measure zero and in the sequel we will use \sup for the essential supremum. The symbol $\sigma(\cdot)$ denotes the maximum singular value of a matrix. The singular values of a matrix A are the non-negative square roots of the eigenvalues of the Hermitian form $A^H A$.

The Fourier transform of a function x is

$$\hat{x}(j\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt \quad (2.5)$$

The operation of the Fourier transform is a linear isometry of L_2^{rxm} onto L_2^{rxm} . The Parseval formula relates inner products

$$x \cdot y = \frac{1}{2\pi} \hat{x} \cdot \hat{y} \quad (2.6)$$

and specializes to

$$\|x\|_2^2 = \frac{1}{2\pi} \|\hat{x}\|_2^2 \quad (2.7)$$

An operator \underline{G} is a mapping which associates with each function in its domain exactly one function in its range. For our purposes, the domain of an operator will be L_{2e}^m and its range will be some subset of L_{2e}^r . The mapping of an operator is denoted

$$y = \underline{G}x \quad (2.8)$$

It is assumed that \underline{G} is a causal operator. Causality means that the output of \underline{G} at time t , does not depend on values of the input at future times, say $t_2 > t_1$.

An operator has a norm induced by the norms on L_{2e}^m and L_{2e}^r . The induced operator norm is

$$\|\underline{G}\|_{2 \rightarrow 2} = \sup_{\tau} \sup_{\|x\|_{2,\tau} = 1} \frac{\|\underline{G}x\|_{2,\tau}}{\|x\|_{2,\tau}} \quad (2.9)$$

The causal operator \underline{G} is said to be L_{2e} stable if it has finite gain, i.e., $\|\underline{G}\|_{2 \rightarrow 2} < \infty$. Thus L_{2e} stable operators map L_2^m into L_2^r .

We now restrict attention to linear operators \underline{G} defined by the convolution integral

$$(\underline{G}x)(t) = \int_{-\infty}^{\infty} g(t-\tau) x(\tau) d\tau \quad (2.10)$$

where g is absolutely integrable. Fubini's theorem provides a relationship between the Fourier transform of $y=Gx$ and the transforms of g and x ,

$$\hat{y} = G \hat{x} \quad (2.11)$$

We refer to G as the transfer matrix of the operator \underline{G} . The Fourier transform converts the original convolution into multiplication in the transformed (frequency) domain. Moreover, the Fourier transform relates the norm of an L_2 stable operator \underline{G} to the transform G . By the Parseval formula (2.7),

$$\|\underline{G}\|_{2 \rightarrow 2} = \sup_{\|x\|_2 \neq 0} \frac{\|\underline{G}x\|_2}{\|x\|_2} = \sup_{\|\hat{x}\|_2 \neq 0} \frac{\|\hat{G}\hat{x}\|_2}{\|\hat{x}\|_2} \quad (2.12)$$

In fact, it can be shown [6] that

$$\|\underline{G}\|_{2 \rightarrow 2} = \|G\|_{\infty} = \sup_{\omega} \bar{\sigma}(G(j\omega)) \quad (2.13)$$

Thus for stable convolutional operators, the norm is simply equal to the supremum over frequency of the largest singular value of the transfer matrix. This makes convolutional operators into a normed linear algebra.

2.2 Performance Measures for Linear Systems

Two alternate measures of performance for linear systems are related to the basic concepts of functional analysis in this section. Broadly speaking, "good" performance means that some error response is "small" in an appropriate sense. An example of such an error response is the classical output sensitivity function of a feedback loop which relates command following errors to output commands. Another example is the response at the input of a plant to sensor noise on the measured variable. We will characterize an error response as an operator and measure its size through norms on the operator. The operator is denoted \underline{G} and maps inputs u in L_2^m into outputs y in L_2^r .

The first type of performance considered is stochastic performance, i.e. the statistical behavior of the error subject to random inputs. If the input u is zero mean white noise with unit intensity, then the covariance of the error is simply

$$E(|y(t)|^2) = \|g\|_2^2 = \frac{1}{2\pi} \|G\|_2^2 \quad (2.14)$$

This performance measure is the integral in the frequency domain of $\text{Tr}(G^H G)$. More generally, the input can be colored noise formed by shaping white noise through the operator \underline{R}^{-1} . Also, we have the freedom to examine the covariance of a filtered version of the error, say $z = \underline{L}y$. For example, such a filter operator could be used to emphasize a particularly crucial band in the frequency domain or to deemphasize low or high frequencies. Including the operators \underline{L} and \underline{R}^{-1} yields

$$E(|z(t)|^2) = \frac{1}{2\pi} \|\underline{LGR}^{-1}\|_2^2 \quad (2.15)$$

The stable operators \underline{L} and \underline{R}^{-1} serve as weightings in the frequency domain in the integral (2.15). This measure of performance has received a great deal of attention in the literature on Wiener and Kalman filtering and the linear-quadratic-Gaussian control problem.

The second type of performance measure is "worst case" performance. The idea in this case is to let the input be any arbitrary function in L_2^m with unit norm. Then the size of the error in the worst case is

$$\sup_{\|u\|_2=1} \|y\|_2 = \|\underline{G}\|_{2 \rightarrow 2} = \|G\|_{\infty} \quad (2.16)$$

Good performance in this sense requires having a small value of $\bar{\sigma}(G(j\omega))$ at every frequency ω . The utility of this approach to measuring performance is increased by introducing weighting operators \underline{L} and \underline{R} much as was done for stochastic performance.

With the weightings, we have

$$\sup_{\|Ru\|_2=1} \|\underline{L}y\|_2 = \|\underline{LGR}^{-1}\|_\infty \quad (2.17)$$

This measure of performance can be written in a slightly different form to emphasize its use as a performance specification. Mathematically,

$$\begin{aligned} \|\underline{L}y\|_2 &< 1 \quad \text{for all } \|Ru\|_2 < 1 \\ \text{iff } \|\underline{LGR}^{-1}\|_{2 \rightarrow 2} &< 1 \\ \text{iff } \|\hat{\underline{LGR}}^{-1}\|_\infty &< 1 \end{aligned} \quad (2.18)$$

This says that having the maximum singular value of the (weighted) transfer matrix less than one at every frequency is necessary and sufficient for the (weighted) error to have norm less than one for any (weighted) input with norm no larger than one. This measure has been the subject of recent theoretical interest within the control community [7], [8]. Also, when viewed as condition on the norm of \underline{LGR}^{-1} , this measure of performance is applicable to nonlinear systems.

Both of these performance measures are expressed in terms of norms on weighted transfer matrices. Both are useful in the analysis of linear systems. The design problem in the first case is to minimize "average" error in the integral square sense. In the second case, the design problem is a mini-max problem: minimize the worst case error in the integral square sense. Note that the weightings on inputs and outputs can always be absorbed into the input-output transfer function. Thus, while weightings are essential to adequately model physical signals, the analysis need not distinguish between weighted or unweighted transfer functions. As an aside, we note that the performance measure in (2.15) can also be interpreted as a mini-max problem with the error signal measured by the L_∞ norm and inputs in L_2 .

2.3 Sectors

In the last several years, the sector concept ([9],[10],[11]) has been recognized as an important tool in feedback design and analysis. The basic idea is that very complicated plant operators (perhaps nonlinear, infinite dimensional, time-varying) can be reliably approximated by simple sector centers (usually finite dimensional, linear, time-invariant systems), provided that the approximation error is properly accounted for in the design process. This "proper accounting" usually means that a design based on the sector center must be restricted to maintain stability. Such restrictions generally increase as the magnitudes of the approximation errors grow. This section provides an introduction to sectors and gives a stability test in terms of a sector condition. It concludes with an interpretation of the stability test as imposing restrictions on the nominal design.

In abstract terms, a sector condition is a functional inequality describing the set of operators in a specified neighborhood of some nominal operator. Formally, the sector $(\underline{C}, \underline{L}, \underline{R})$ is the set of all operators \underline{G} mapping L_{2e}^m into L_{2e}^r which satisfy

$$\|\underline{L}(\underline{G}-\underline{C})\|_{2,\tau} \leq \|\underline{R}x\|_{2,\tau} \quad \forall \tau \in \mathbb{R}_+, x \in L_{2e}^m \quad (2.19)$$

The nominal operator \underline{C} is referred to as the center of the sector. The operators \underline{L} and \underline{R} specify the size of the neighborhood about \underline{C} . We will require \underline{L} and \underline{R} to be L_{2e} -stable operators and to have L_{2e} -stable inverses. When \underline{L} and \underline{R} are linear, time-invariant operators, this means they have no poles or zeros in the right-half of the complex plane and no excess of poles or zeros. In the case of an L_{2e} -stable center, condition (2.19) can be rewritten as

$$\| \underline{L}(\underline{G}-\underline{C})\underline{R}^{-1} \|_{2 \rightarrow 2} \leq 1 \quad (2.20)$$

It is noteworthy to compare this condition with the performance specification (2.18).

2.4 Stability Conditions

Having introduced sectors, we now turn to characterizing the stability of a feedback system when an operator in the system is described by a sector. Consider the feedback configuration shown in Figure 2-1. The two closed-loop operators of this system are E_1 and E_2 , mapping both inputs u_1 and u_2 into the outputs e_1 and e_2 , respectively. It is assumed that the system is causal and well-posed, [6] i.e., the operators \underline{M} , \underline{D} , E_1 , and E_2 are all causal. The system is stable if both E_1 and E_2 are.

Theorem 2-1. (Small Gain Theorem) Under these conditions, the closed-loop system is L_{2e} -stable if

$$\| \underline{M} \|_{2 \rightarrow 2} \| \underline{\Delta} \|_{2 \rightarrow 2} < 1$$

A brief discussion of the proof of this theorem follows. The error e_1 is given by

$$\begin{aligned} e_1 &= u_1 - \underline{\Delta} e_2 \\ &= u_1 - \underline{\Delta} u_2 - \underline{\Delta} \underline{M} e_1 \end{aligned} \quad (2.24)$$

and so

$$(I + \underline{\Delta} \underline{M}) e_1 = u_1 - \underline{\Delta} u_2 \quad (2.25)$$

The norm of the right-hand side of (2.25) is bounded from above by

$$\| u_1 \|_2 + \| \underline{\Delta} \|_{2 \rightarrow 2} \| u_2 \|_2$$

and the left-hand side is bounded from below by

$$(1 - \| \underline{\Delta} \|_{2 \rightarrow 2} \| \underline{M} \|_{2 \rightarrow 2}) \| e_1 \|_2$$

Combining these bounds and using the inequality of the theorem yields

$$\| e_1 \|_2 \leq \frac{1}{1 - \| \underline{M} \|_{2 \rightarrow 2} \| \underline{W} \|_{2 \rightarrow 2}} (\| u_1 \|_2 + \| \underline{W} \|_{2 \rightarrow 2} \| u_2 \|_2) \quad (2.26)$$

Thus the operator E_1 has finite gain and is stable. Stability of E_2 is shown similarly.

The Small Gain Theorem will now be used to obtain stability robustness conditions for a feedback system containing a sector. Stability robustness conditions will be obtained which guarantee the stability of the closed loop system for any operator that is an element of the sector. The feedback system under consideration is shown in Figure 2-2. Also shown in the figure are two alternative representations of the feedback loop. The first alternative uses the earlier observation that a sector may be expressed as the parallel combination of its center and a sector with the same radius centered about the null operator. The second alternative representation employs the closed loop operator for the feedback combination of \underline{G} and \underline{C} . It is assumed that the closed loop operator \underline{M} is L_{2e} -stable.

The Small Gain Theorem will be applied to the feedback system in Figure 2-2 involving operator \underline{M} and sector $(\underline{C}, \underline{L}, \underline{R})$. Let the operator $\underline{\Delta}$ be any element $(\underline{C}, \underline{L}, \underline{R})$. By (2.20), this means that

$$\| \underline{L} \underline{\Delta} \underline{R}^{-1} \|_{2 \rightarrow 2} < 1 \quad (2.27)$$

Rather than applying the Small Gain Theorem directly to \underline{M} and $\underline{\Delta}$, we apply it to the operators $\underline{R} \underline{M} \underline{L}^{-1}$ and $\underline{L} \underline{\Delta} \underline{R}^{-1}$, shown in Figure 2-3. These operators are both L_{2e} -stable, and closed loop stability of the system in Figure 2-3 is equivalent to stability of the representations in Figure 2-2. By the Small Gain Theorem, closed loop stability is guaranteed if

$$\|RML^{-1}\|_{2 \times 2} \|LAR^{-1}\|_{2 \times 2} < 1 \quad (2.28)$$

When combined with (2.27),

$$\|RML^{-1}\|_{2 \times 2} \leq 1 \quad (2.29)$$

suffices to ensure closed loop stability. This is summarized in the following theorem.

Theorem 2-2. Consider the feedback system of Figure 2-2 and assume $(C+G^{-1})^{-1}$ is L_2 -stable. If

$$\|R(C+G^{-1})^{-1}L^{-1}\|_{2 \times 2} \leq 1 \quad (2.30)$$

then the closed loop system is stable for any operator in the sector (C, L, R) .

As in the case of weightings on inputs and outputs, the weightings and center that determine a sector can be absorbed into the interconnecting transfer function. Thus, without loss of generality, we can restrict attention to sectors of the form sector (O, I, I) . In this case (2.30) becomes

$$\|G\|_{2 \times 2} \leq 1 \quad (2.31)$$

and for LTI G ,

$$\|G\|_{\infty} \leq 1 \quad (2.32)$$

It will be helpful to compare the stability robustness condition of Theorem 2.2 with the performance specification (2.18). Stability robustness is ensured if a closed loop operator when "weighted" by R and L^{-1} has norm less than one. Performance is achieved if some closed loop error operator has "weighted" norm less than one. The same type of condition applies in both cases. Because the same condition is used for both a performance specification and stability robustness, we will be able to obtain a combined condition which can guarantee simultaneous robust performance and stability.

3. ANALYSIS

The mathematical tools of the previous section can be put together to analyze the performance and robustness properties of control systems. The remainder of this paper explores various modelling assumptions and the impact these assumptions have on analysis methods. This section summarizes these assumptions and the resulting analysis tools. Referring to Figure 3-1a, the nominal model is assumed throughout to be a Linear, Time-Invariant Ordinary Differential Equation (LTIODE). The uncertain inputs are assumed to be either filtered white noise or weighted L_2 -norm bounded signals. The plant uncertainty is modelled as perturbations (not necessarily small) to the nominal. Performance is measured in terms of either the weighted error covariance or the weighted L_2 -norm of the error.

These performance measures are intended to reflect engineering issues such as good command response or small errors in regulation or estimation. Perturbations typically arise in an attempt to model changes and uncertainty in operating conditions and plant characteristics as well as unmodelled dynamics. Uncertain inputs model disturbances, noises and commands. The analysis framework used in this paper includes all the standard linear time-invariant filtering and control problems, including the so-called two-degree-of-freedom control problem. This last problem is obtained when commands are modelled in the usual way as uncertain input signals.

Since the focus of this section is on analysis, the controller can be viewed as just another system component. Thus for analysis purposes, Figure 3-1a may be reduced to Figure 3-1b. Here P is a 2×2 block transfer function matrix providing connections from external inputs and perturbations to outputs and perturbations. Without loss of generality, the input signals, errors and perturbations can be left unweighted since any weightings can be absorbed into the interconnection operator P . Note that any

interconnection of inputs and outputs with components and perturbations, including weightings, may be rearranged into this form. Then the output can be written as

$$e = (P_{22} + P_{21} W(I - P_{11} W)^{-1} P_{12})u. \quad (3.1)$$

It is assumed that stability is always a performance requirement and thus that P has all its poles in the open left-half plane.

The standard modelling assumptions and the resulting analysis methods are summarized in Table 3-1. The first option is that uncertainty is modelled as white noise and performance is measured in terms of error covariance. From (2.14), the error covariance can be evaluated in terms of the L_2 -norm of P_{22} . This model is appealing in that many physical noises and disturbances have existing, accepted models as filtered white noise and that computation of $\|P_{22}\|_2$ is quite easy using Lyapunov equations. Furthermore, synthesis (the so-called Wiener-Hopf-Kalman (WHK) theory) in this context involves linear approximation in a Hilbert space, also computationally appealing. Unfortunately, few physical systems are adequately modelled with additive white noise as the only uncertainty.

Some alternatives to the white noise view of uncertainty are summarized in the remainder of Table 3-1. These options may be thought of as being separate cases of the following general form of a performance/robustness theorem:

Given	Modelling Assumptions	
Performance Requirement	iff	Analysis Test
is satisfied		is satisfied

(3.2)

The first alternative (Case 2) involves modelling inputs as unknown-but-bounded (in an L_2 -sense) and requiring that the output remain bounded (in L_2) below a specified level for all such inputs. From (2.19) the resulting analysis test involves the L_∞ -norm of P_{22} , the same norm which appears in the robust stability test of Case 3 (from 2.32). The chief advantage of these assumptions over Case 1 is that both uncertain inputs and plant perturbations are handled with the same $\|\cdot\|_\infty$ test. A less compelling reason is that designs for unknown-but-bounded L_2 inputs and outputs can be given a mini-max energy interpretation.

Note that any induced operator norm would provide an analysis test that would handle both uncertain norm-bounded inputs and (induced) norm-bounded perturbations. For example, modeling signals as unknown-but-bounded in magnitude (and using the resulting induced norm for convolution operators on L_∞ for analysis) has obvious advantages in applications where signal magnitude is a more natural notion than energy. On the other hand, there are some important reasons for choosing L_2 signal models, $\|\cdot\|_\infty$ perturbation bounds, and $\|\cdot\|_\infty$ analysis tests over, say, other L_p spaces:

1. The induced convolution-operator norm on L_2 (i.e., $\|\cdot\|_\infty$ on transfer functions) is the only induced norm which yields necessary as well as sufficient robust stability tests.
2. Perturbation models of this type are currently the most easily obtained.
3. An optimal synthesis theory analogous to that of WHK is now available (i.e., the main result of reference [8]).

4. Engineers have developed substantial experience with these methods through the use of Bode plots and more recently, their singular value generalizations [1].

Clearly, these reasons are not entirely independent. The $\|\cdot\|_\infty$ norm on transfer functions is reasonably easily computed, but it does involve a search over one frequency variable.

It should be reiterated that in practice the use of weights on signals and perturbations is essential, since both vary with direction and frequency. This is true independent of the particular assumptions being made. By absorbing any weightings into the interconnection function P , the weighted case can be reduced to that considered in Figure 3-1 and Table 3-1. This is one advantage of the framework proposed here over less general ones in that any interconnection of signals, systems and perturbations, including weights, can be rearranged to fit the framework.

While Cases 2 and 3 provide a single framework in which to analyze performance and robustness (of stability), the $\|\cdot\|_\infty$ norm alone provides no systematic, reliable method for analyzing robust performance. Furthermore, $\|\cdot\|_\infty$ analyzes robustness with respect to purely unstructured uncertainty. A more sophisticated tool that treats robust performance with respect to structured uncertainty involves the structured singular value, μ , and $\|\cdot\|_\mu$. This is the subject of the remainder of the paper.

Cases 4 and 5 of Table 3-1 summarize the two basic applications of μ to analysis. Case 4 gives a structured version of Case 3 by characterizing robust stability with respect to block-diagonal perturbations. This is quite general since any interconnection of perturbations can be rearranged to fit the structure of Figure 3-1 with a block diagonal Δ . Case 5 generalizes cases 2, 3, and 4 by characterizing the performance (in an L_2 -bounded sense) for systems with structured uncertainty. This is currently the only available method for systematically analyzing the performance of complex systems with plant perturbations. Many aspects of LSS control problems naturally fit in the framework outlined in this section. Table 3-2 lists some common sources of uncertainty in the form of commands, disturbances and noises, and perturbations. Figure 3-2 shows schematically how uncertain models for actuators, sensors and the spacecraft structure can be rearranged to fit the framework of Figure 3-1. The remainder of the paper will review the details of analysis using $\|\cdot\|_\mu$.

4. FEEDBACK ANALYSIS AS A BLOCK-DIAGONAL BOUNDED PERTURBATION PROBLEM

4.1 Robustness Characterization

This section formulates the basic feedback problem of achieving performance in the face of uncertainties as a stability problem in the presence of block-diagonal bounded perturbations [4]. The formulation involves sector-bounded transfer functions as basic building blocks. The robustness and performance properties of a feedback system will be expressed in terms of a collection of linear, time-invariant operators, Δ_i , $i=1, 2, \dots, m$, each of which is an element of a sector with zero center, i.e., $\Delta_i \in (0, L_i, R_i)$. These operators are the basic building blocks in a combined robustness/performance characterization of feedback systems.

The use of sector bounded, linear, time-invariant operators to characterize robustness has been a central theme in many recent references, including [1] where such operators were inserted at the inputs or outputs of a plant model in order to represent so called unstructured uncertainties (modeling errors with no assumed structure except for known magnitude bounds on their transfer functions). Necessary and

sufficient conditions such as Theorem 2-2 were then derived for stability robustness in the face of such uncertainties. For example, a stable feedback loop with plant G and compensator K will remain stable in the face of all possible perturbed plants $G' = [I + \Delta]G$, with $\Delta \in (0, L, R)$, if and only if

$$\| RGK(I+GK)^{-1} L^{-1} \|_{\infty} \leq 1 \quad (4.1)$$

Note that with R and L specified, this inequality imposes conditions on the shape of the closed loop frequency response, $GK(I + GK)^{-1}$, which must be satisfied in order to assure robust stability. These conditions are unique to the assumed form of plant perturbations (e.g., $G' = (I + \Delta)G$ in the present case). Each such assumed form corresponds to a specific location where Δ is inserted in the nominal feedback loop. The location for our present case is shown in Row 1 of Table 4-1. Other locations correspond to other assumed forms for G' and produce different necessary and sufficient stability robustness conditions. A representative set of possibilities is summarized in the remaining rows of Table 4-1. Of course, each case in Table 4-1 is a special case of Figure 3-1.

Table 4-1 also indicates representative types of physical uncertainties which can be usefully represented by sector bounded perturbations inserted at the indicated locations. For example, the representation $G' = (I + \Delta)G$ in Row 1 is useful for output errors at high frequencies, covering such things as unmodelled high frequency dynamics of sensors or plant, including diffusion processes, transport lags, electro-mechanical resonances, etc. The representation $G' = G(I + \Delta)$ in Row 2 covers similar types of errors occurring at the inputs. Both cases should be contrasted with Rows 4 and 5 which treat $G' = (I + \Delta)^{-1}G$ and $G' = G(I + \Delta)^{-1}$. These representations are more useful for variations in modelled dynamics, such as low frequency errors produced by parameter variations with operating conditions, with aging, or across production copies of the same plant. Discussion of still other cases is left to the table. Note from the table that the stability requirements on Δ do not limit our ability to represent variations in either the number or location of rhp singularities.

The most significant thing to understand about Table 4-1 is that the stability robustness conditions shown are sufficient to ensure stability only if all the uncertainties occur at the indicated locations and none occur elsewhere. In order to use the conditions directly, therefore, designers are obliged to reflect all known sources of uncertainty from their known point of occurrence to a single reference location in the loop. Such reflected uncertainties invariably have a great deal of structure which must then be "covered up" with a larger, arbitrarily more conservative perturbation in order to maintain a simple cone bounded representation at the reference location. By "arbitrarily more conservative," we mean that examples can be constructed where the degree of conservatism is arbitrarily large. Of course, other examples exist where it is quite reasonable.

Alternatively, designers could choose to treat uncertainties occurring at several different locations in the feedback loop as a single uncertainty occurring at one location in a larger feedback loop. To be specific about this alternative, let Δ_i , $i=1, 2, \dots, m$, denote a collection of such uncertainties positioned at location l_i , $i=1, 2, \dots, m$. Note that at each l_i , the feedback loop has an input, where it receives the signals from Δ_i , and also an output, where it supplies signals to Δ_i . Let M_{ii} be the operator between these two sets of signals. Further, let M_{ij} denote the operator between the inputs at location l_j and the outputs at location l_i . Then the block-structured operator

$$M = \{ M_{ij}^{\Delta} \} \quad (4.2)$$

represents all interactions of the feedback loop with its uncertainties, and indeed, the block-diagonal bounded perturbation diagram in Figure 4-1 is an equivalent representation of the loop. Here we have $\Delta = \text{diag}(\Delta_1, \Delta_2, \dots, \Delta_m)$.

Note that the feedback elements in this larger loop are zero in the absence of uncertainties. Hence, M will be a stable "plant" whenever the original nominal loop is stable. As an example of this representation, consider the system in Figure 4-2. This system, with two uncertainties present simultaneously, the first from Row 2 and the second from Row 4 of Table 4-1, is described by the following M operator:

$$M = \begin{pmatrix} (I + KG)^{-1}KG & (I + KG)^{-1}K \\ (I + GK)^{-1}G & (I + GK)^{-1} \end{pmatrix} \quad (4.3)$$

Given the equivalent system in Figure 4-1 with sector bounded Δ_1 , it follows from the Small Gain Theorem that the loop remains stable in the presence of these uncertainties if

$$\|RML^{-1}\|_2 \leq 1$$

or, in the frequency domain,

$$\bar{\sigma} [R(j\omega)M(j\omega)L(j\omega)^{-1}] \leq 1 \quad \forall \omega \quad (4.4)$$

where $R = \text{diag}(R_1, R_2, \dots, R_m)$ and $L = \text{diag}(L_1, L_2, \dots, L_m)$. This condition provides an alternate test for stability robustness. Like the procedure of reflecting all uncertainties to one reference location, however, the new test can be arbitrarily more conservative because it ignores the known block-diagonal structure of the uncertainties in Figure 4-1.

One of the objectives of the results in this paper is precisely to reduce the conservatism of robustness and performance tests for block-diagonal structures such as Figure 4-1. We do this by introducing a generalized notion of the maximum singular value for block-diagonal structures. This generalization is developed in Section 5. It is called the structured singular value (SSV) and is denoted by the symbol μ . It yields the following necessary and sufficient conditions for robust stability of the BDBP problem:

$$\mu[R(j\omega)M(j\omega)L^{-1}(j\omega)] \leq 1 \quad \forall \omega \quad (4.5)$$

This represents our extension of the Small Gain Theorem which we call the Small μ Theorem.

Since all simultaneous uncertainties can be put into block-diagonal form by merely constructing the associated operator RML^{-1} , the SSV allows us to nonconservatively analyze simultaneous occurrences of uncertainties anywhere in a feedback system. The uncertainties may be sector bounded errors of individual components of the system (SISO or MIMO), they may be individual parameter variations in the model, or even polynomial approximations of parameters entering nonlinearly. In fact, the only restriction which remains is that all variations must be allowed to be complex. Pure real variations or pure imaginary variations cannot be separated into individual blocks.

4.2 Performance Characterization

The ability to treat simultaneous, structured uncertainties also offers, almost as a free byproduct, the ability to deal simultaneously with the performance and robustness aspects of feedback. As discussed in 2.2 and 2.4, the test for satisfaction of a performance specification is identical to a test for robustness with respect to some uncertainty. That is, any performance specification has a corresponding robustness requirement such that one is satisfied if and only if the other is. Thus the SSV tests for robustness can be used directly to evaluate performance.

The equivalence between performance and robustness is elaborated in Column 4 of Table 4-1, where each of the conditions imposed on feedback loop shapes by perturbation Δ_1 at location \mathbf{L}_1 is given a performance interpretation. For example, the perturbations in Row 4 impose requirements (through L and R) on the operator $(I + KG)^{-1}$. This operator is, of course, the classical (output) sensitivity function of the feedback loop. Small values over some frequency range guarantee low closed loop sensitivity to open loop variations and low command following errors to output commands over that range. The test for satisfaction of any L_2 -performance specification is identical to test for robustness with respect to some perturbation. That is, any performance specification has a corresponding robustness requirement such that one is satisfied if and only if the other is. Thus, the structured singular value tests for robustness can be used directly to evaluate performance. This may be thought of as introducing a "fictitious uncertainty" to impose a performance requirement.

To illustrate how such fictitious uncertainties actually enforce performance specs, consider the simple case where a single true uncertainty, say Δ_r from Row 2, and a single fictitious (performance) uncertainty, say Δ_p from Row 4, are specified for our feedback system. Let the structured singular value condition (4.5) be satisfied for the corresponding M matrix, (4.3). Then the system remains stable in the face of Δ_r and Δ_p occurring simultaneously. Obviously, it will also remain stable for Δ_p with $\Delta_r = 0$. This means that the nominal system must satisfy the performance condition

$$\bar{\sigma}[R_p(I + KG)^{-1}L_p^{-1}] \leq 1 \quad \forall \omega \quad (4.6)$$

because the latter is also a necessary and sufficient condition for robust stability with Δ_p only. This much is straightforward. What is not so evident but much more important is that condition (4.6) is also satisfied for all perturbed feedback loops. That is, for all true plants $G' = G(I + \Delta_r)$ we have

$$\bar{\sigma}[R_p(I + KG')^{-1}L_p^{-1}] \leq 1 \quad \forall \omega \quad (4.7)$$

Hence, the performance spec is satisfied in the face of all possible true uncertainties. A proof of this consequence of the structured singular value condition is discussed in Section 5.

5. STRUCTURED SINGULAR VALUE ANALYSIS OF FEEDBACK SYSTEMS

5.1 Introduction to the Structured Singular Value

We have discussed how the problem of analyzing performance in the face of structured uncertainty can be expressed as a BDBP problem. The standard singular value tests applied to the BDBP can be excessively conservative because they ignore the block-diagonal structure. A more general nonconservative test (the Small μ Theorem) is developed in this section which removes this limitation. By nonconservative we mean providing a necessary and sufficient condition. The test is expressed in terms of a new measure, the structured singular value μ . This section begins with review of the results in [5] where μ was introduced.

To provide a more precise description of block-diagonal perturbations, let $K = (m_1, m_2, \dots, m_n, k_1, k_2, \dots, k_n)$ be a $2n$ -tuple of positive integers. All the definitions that follow depend on K, but to simplify the notation this dependency will not be explicitly represented. Let

$$k = \sum_{j=1}^n m_j k_j \quad \text{and} \quad m = \sum_{j=1}^n m_j$$

Let X be a set of $k \times k$, rational, block-diagonal matrices defined by

$$X = \{\text{diag}(\underbrace{\Delta_1, \Delta_1, \dots, \Delta_1}_{m_1}, \underbrace{\Delta_2, \Delta_2, \dots, \Delta_2}_{m_2}, \dots, \underbrace{\Delta_n, \Delta_n, \dots, \Delta_n}_{m_n}) \mid \text{for each } j=1, 2, \dots, n, \Delta_j \text{ is a } k_j \times k_j \text{ matrix}\} \quad (5.1)$$

Let U be the set of block-diagonal unitary matrices, and \mathcal{B} the set of real diagonal matrices such that

$$\mathcal{B} = \{\text{diag}(d_1 I_{k_1}, d_2 I_{k_2}, \dots, d_m I_{k_m}) \mid d_i \in \mathbb{R}^+ (0, \infty)\} \quad (5.2)$$

What is desired is a function (depending on K)

$$\mu: (k) \rightarrow [0, \infty) \quad (5.3)$$

with the property that $\forall M \in (k)$

$$\begin{aligned} \det(I + M\Delta) \neq 0 \quad \forall \Delta \in X, \quad \sigma(\Delta) \leq \delta \\ \text{iff } \delta \mu(M) \leq 1 \end{aligned} \quad (5.4)$$

This could be taken as a definition of μ . Alternatively, μ could be defined as

$$\begin{aligned} \mu(M) = & \begin{cases} 0 & \text{if no } \Delta \in X \text{ solves } \det(I + M\Delta) = 0 \\ \min_{\Delta \in X} \{\bar{\sigma}(\Delta) \mid \det(I + M\Delta) = 0\}^{-1} & \text{otherwise} \end{cases} \end{aligned} \quad (5.5)$$

This definition shows that a well-defined function satisfies (5.4). It probably has little additional value since the optimization problem involved does not appear to have useful properties.

Using these definitions, the following useful properties of μ are easily proven.

- | | |
|---|---|
| 1) $\mu(aM) = a \mu(M)$ | $\forall M \in \mathcal{M}(k)$ |
| 2) $\mu(I) = 1$ | |
| 3) $\mu(AB) \leq \sigma(A) \mu(B)$ | $\forall A \in X, B \in \mathcal{M}(k)$ |
| 4) $\mu(\Delta) = \sigma(\Delta)$ | $\forall \Delta \in X$ |
| 5) If $n=1$ and $m_1=1$ then $\mu(M) = \sigma(M)$ | $\forall M \in \mathcal{M}(k)$ |
| 6) If $n=1, k_1=1$, then $k=m_1$,
$X = \{\lambda I \mid \lambda \in \mathbb{R}\}$ and $\mu(M) = \rho(M)$ | $\forall M \in \mathcal{M}(k)$ |
| 7) If $\Delta \in X, U \in \mathcal{U}$ then $U\Delta U^H$ and $\Delta U \in X$ and $\sigma(U\Delta) = \sigma(\Delta U) = \sigma(\Delta)$ | |
| 8) $\forall \Delta \in X$ and $\forall D \in \mathcal{B}$ | $D\Delta D^{-1} = \Delta$ |
| 9) $\forall U \in \mathcal{U}$ and M | $\mu(MU) = \mu(UM) = \mu(M)$ |
| 10) $\forall D \in \mathcal{B}$ and M | $\mu(DMD^{-1}) = \mu(M)$ |
| 11) $\max_{U \in \mathcal{U}} \rho(UM) \leq \mu(M) \leq \inf_{D \in \mathcal{B}} \bar{\sigma}(DMD^{-1})$ | $\forall M \in \mathcal{M}(k)$ |

Properties 5 and 6 show that the structured singular value has, as special cases, both the spectral radius and the maximum singular value. Property 9 means that μ is U -invariant.

The most important results from [5] are the following, which deal with the bounds in property 11:

- A) The left-hand-side inequality in property 11 is always an equality. This expresses μ in familiar linear algebraic terms, but the optimization problem involved may have multiple local maxima.

- B) The right-hand-side inequality in 11 is an equality when there are three or fewer blocks, and the blocks are not repeated. The blocks themselves, and therefore M , may be of arbitrarily large dimension. A tedious but straightforward computation shows that the optimization problem involved is always convex [12]. Furthermore, the minimization is over only $n-1$ parameters for n blocks, independent of block size, making this an attractive alternative to A.

Note that the transformation DMD^{-1} is simply a rescaling of the inputs and outputs of M . The SSV is invariant with respect to such rescaling (property 10), while singular values do, of course, vary with rescaling. This implies, for example, that the ad hoc method of performing a change of units can reduce the conservatism associated with singular values. For some time we have been using Osborne's technique [13], which minimizes the Frobenius norm of DMD^{-1} to compute frequency-dependent D matrices. We now have new algorithms which compute D to directly minimize $\sigma(DMD^{-1})$.

While the strongest results on the use of the DMD^{-1} scaling are for nonrepeated blocks, the use of scaling before computing singular values can be quite effective in treating repeated block problems. These typically arise when analyzing robustness with respect to variations in scalar parameters occurring in several places within a system. This leads naturally to problems where the block-diagonal perturbations are made of either nonrepeated matrix blocks or repeated scalars. In this case the set x takes the special form:

$$x = \{\text{diag}(\lambda, I, \lambda_2 I, \dots, \lambda_{n_1} I, \Delta_1, \Delta_2, \dots, \Delta_{n_2})\}$$

Suppose now we let

$$D = \{\text{diag}(D_1, D_2, \dots, D_{n_1}, d_1 I, d_2 I, \dots, d_{n_2} I)\} \quad (5.6)$$

where the structure of D matches x and the D_i are full block matrices and the d_i are real scalars. Then an upper bound for μ with this structure is again

$$\mu(M) \leq \inf_{D \in \mathcal{S}} \bar{\sigma}(DMD^{-1})$$

The use of full block D_i 's improves the bound and a simple argument shows that the problem is still convex. Of course, the scaling could be extended to handle repeated nonscalar blocks, but these have not yet been found to be applicable to system problems.

Numerical software for computing μ has been developed using algorithms based on these results. In addition to using this software to analyze some simple feedback designs, test runs have been made on a large number of psuedo-random matrices. It appears that the global maximum in A is often easily found, although a simple gradient search is inadequate. Also, the bound obtained in B appears to be quite good (to within 15%) for cases of more than 3 blocks. These observations are most encouraging, especially considering the experimental and preliminary nature of the software.

There are essentially two direct applications of singular values to the BDBP problem, which provide bounds for μ :

- 1) Ignore the block-diagonal structure and compute $\sigma(M)$. This gives an upper bound for μ .
- 2) Treat each perturbation one at a time. Compute the largest maximum singular value for each of the corresponding diagonal blocks. This gives a lower bound for μ .

The gap between these two bounds may be arbitrarily large.

5.2 Robustness Analysis: The Small μ Theorem

The preceding discussion of μ and the BDBP problem has dealt with determining the size of the minimum structured perturbation Δ that causes $I + M\Delta$ to be nonsingular. We are interested in using the structured singular value to answer robustness, sensitivity, and performance questions for multivariable feedback systems. The connection between μ and these essential feedback properties is provided by the Small μ Theorem, which characterizes the stability robustness properties of a feedback system with respect to block-diagonal perturbations. In order to state the Small μ Theorem we need the following additional definitions depending on K :

$$X_r = \{ \Delta c H_\infty \mid \Delta(j\omega) c X V \omega \} \quad (5.7)$$

Thus X_r is the set of all H_∞ transfer functions which are block-diagonal with the structure defined by K . Let $BX_r = BH_\infty X_r$.

Now, consider the system in Figure 3-1 and equation (3.1) with $\Delta c BX_r$. This is Case 4 of Table 3-1. Define stability in this case to be equivalent to $I - P_{11}\Delta$ nonsingular in the closed right-half-plane. This is stronger than bounded-input-bounded-output stability, but the following theorems apply equally well to either notion of stability. Although μ is not a norm, abuse notation and let

$$\|M\|_\mu = \sup_\omega \mu(M(j\omega))$$

Theorem 5.1 (Small μ): The system in (3.1) is stable for all $\Delta c BX_r$, iff

$$\|P_{11}\|_\mu \leq 1 \quad (5.8)$$

Proof: To prove the if part, suppose $\mu_c < 1$ and let $\Delta c BX_r$, then using properties 3 and 11 and the definition of X

$$\sup_{\text{Res} \geq 0} \rho(P_{11}\Delta) = \sup_{s=j\omega} \rho(P_{11}\Delta) \leq \sup_{s=j\omega} \mu(P_{11}^{-1}) = \mu_c < 1.$$

Thus $I + P_{11}\Delta$ is nonsingular for all $\text{Res} \geq 0$. Since Δ was arbitrary, the canonical system is stable for all $\Delta c BX_r$.

Conversely, suppose $\mu(P_{11}) \mid_{\omega=\omega_0} \geq 1$ (ω_0 may be ∞). Then $\exists \Delta c BX_r$ such that

$\det(I + P_{11}\Delta) \mid_{\omega=\omega_0} = 0$. Thus, $\exists \Delta c BX_r, \exists, \det(I + M\Delta)_{\omega=\omega_0} = 0$ and the canonical system is not stable for all $\Delta c BX_r$. \square

This theorem guarantees that if $m(P_{11})$ is less than 1 at every frequency, then the system is stable for all structured perturbations WcX . Conversely, if $m(P_{11})$ is greater than or equal to 1 at some frequency, then there exists a structured perturbation WcX_r that results in instability.

5.3 Performance Implications

As noted in Section 4, the Small m Theorem can also guarantee a pre-specified performance level by including a performance block in the BDBP problem. Furthermore, this performance level is guaranteed for all structured perturbations WcX . These claims are made precise by a corollary to the Small μ Theorem that treats performance.

Consider the system in Figure 3-2 and equation (3.1) with $\Delta c BX_r$. Suppose a performance specification is given as

$$\|W\|_2 \leq 1 \text{ for all } \|u\|_2 \leq 1 \text{ and all } \Delta c BX_r$$

From (2.18) this is equivalent to requiring that

$$\sup_{\Delta \in BX_r} \|P_{22} + P_{21} \Delta (I - P_{11} \Delta)^{-1} P_{12}\|_{\infty} \leq 1 \quad (5.9)$$

This is Case 5 of Table 3-1.

Define

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

$$K_T = (m_1, m_2, \dots, m_n, l, k_1, k_2, \dots, k_n, p)$$

$$X_T = \{\Delta = \text{diag}(\Delta_1, \Delta_2) \mid \Delta_1 \in X_r, \Delta_2 \in RH_{\infty}^{p \times p}\}$$

$$BX_T = X_T \cap BH_{\infty}$$

where p is the dimension of P_{22} . For the remainder of this section, let μ be defined in terms of K_T . Then the following theorem gives the relationship between robust performance and the SSV.

Theorem 5.2 (Robust Performance)

The performance requirement (5.9) is satisfied

$$\text{iff} \quad \|P\|_{\mu} \leq 1$$

Proof: Repeated application of the Small μ Theorem yields

$$\mu(P(j\omega)) \leq 1 \quad \forall \omega$$

$$\text{iff} \quad \det(I - P\Delta) > 0 \quad \forall \text{Res} > 0 \quad \forall \Delta \in X_T$$

$$\text{iff} \quad \det(I - P_{11}\Delta_1) \det(I - (P_{22} + P_{12}\Delta_1(I - P_{11}\Delta_1)^{-1}P_{21})\Delta_2) > 0$$

$$\forall \text{Res} \geq 0 \quad \forall \Delta_1 \in BX_r, \Delta_2 \in BH_{\infty}^{p \times p}$$

$$\text{iff} \quad \bar{\sigma}(P_{22} + P_{21}\Delta(I - P_{11}\Delta)^{-1}P_{12}) \leq 1 \quad \forall \text{Res} \geq 0 \quad \forall \Delta \in BX_r$$

$$\text{iff} \quad \sup_{\Delta \in BX_r} \|P_{22} + P_{12}\Delta(I - P_{11}\Delta)^{-1}P_{21}\|_{\infty} \leq 1$$

We note that this theorem extends the Small μ Theorem's robust stability results to a composite, simultaneous result on robust stability and performance. Thus, given an uncertain plant model with structured perturbations and a performance specification, we have a necessary and sufficient condition in terms of μ for satisfaction of the performance spec in the face of the uncertainty. If the condition $\mu \leq 1$ is met, then the desired performance is achieved for all perturbed plants. If $\mu \geq 1$, then there exists a structured perturbation which causes the performance spec to be violated. The robust performance condition may be thought of as arising from an equivalent "fictitious uncertainty," although this interpretation is not necessary.

6. CONCLUDING COMMENTS

This paper outlined an analysis technique based on the Structured Singular Value μ for linear feedback systems that provide a reliable, nonconservative measure of performance in the face of structured uncertainty. The Small μ Theorem gives a necessary and sufficient condition in terms of μ for stability of a linear system with multiple, simultaneous, norm-bounded perturbations of fixed but arbitrary structure. The Robust Performance Theorem provides a similar condition for the satisfaction of performance specifications in the presence of structured perturbations. These results have great potential for application to Large Space Structure Control Problems. Further research is in progress to realize this potential.

References

- [1] J.C. Doyle and G. Stein, "Multivariable Feedback Design: Concepts for a classical/modern Synthesis, IEEE Trans. on Automatic Control, Vol. AC-26, No. 1, February 1981.
- [2] I.M. Horowitz, Synthesis of Feedback Systems, New York: Academic, 1963.
- [3] Special Issue on Linear Multivariable Control Systems, IEEE Trans. on Automatic Control, February 1981.
- [4] J.C. Doyle, J.E. Wall and G. Stein, "Performance and Robustness Analysis for Structured Uncertainty", IEEE Conference on Decision and Control, December 1982.
- [5] J.C. Doyle, "Analysis of Feedback Systems With Structured Uncertainties, "Proc. IEE", November 1982.
- [6] C.A. Desoer and M. Vidyasagar, Feedback Systems: Input-Output Properties, Academic Press, 1975.
- [7] B.A. Francis and G. Zames, "Design of H^∞ - Optimal Multivariable Feedback Systems," IEEE Conference on Decision and Control, December 1983.
- [8] J.C. Doyle, "Synthesis of Robust Controllers and Filters," IEEE Conference on Decision and Control, December 1983.
- [9] G. Zames, "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems - Part I," IEEE Trans. on Automatic Control, Volume AC-11, No. 2 pp. 228-238, April 1966.
- [10] G. Zames, "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems - Part II," IEEE Trans. on Automatic Control, Volume AC-11, No. 3, pp. 465-476, July 1966.
- [11] M.G. Safonov, Stability and Robustness of Multivariable Systems, Cambridge, MA, MIT Press, 1980.
- [12] J.C. Doyle and M.G. Safonov, "Convexity of the Block Diagonal Scaling Problem," Honeywell Internal Memo, August 1982.
- [13] E.E. Osborne, "On Preconditioning of Matrices," I. Assoc. Comput., Mach., Vol. 7, pp. 338-345, 1960.

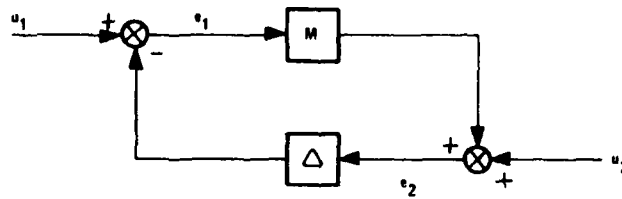


Figure 2-1. Feedback Configuration for the Small Gain Theorem

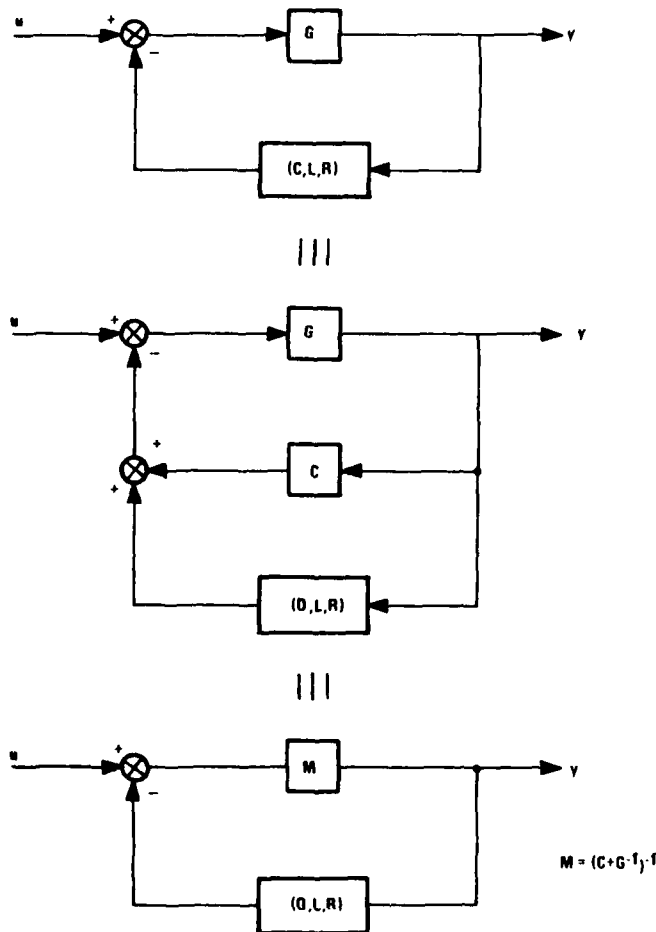


Figure 2-2. Three Equivalent Representations of the Feedback Configuration Involving Operator G and Sector (C, L, R)

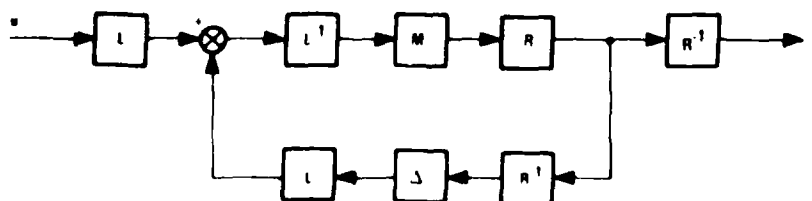
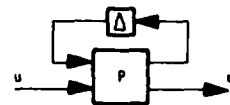


Figure 2-3. Operators L and R with Inverses Do Not Alter Stability of Closed Loop System



$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

$$e = (p_{22} + p_{21} \Delta (I - p_{11} \Delta)^{-1} p_{12}) u$$

CASE	MODELLING ASSUMPTIONS		PERFORMANCE REQUIREMENT	ANALYSIS TEST	SYNTHESIS METHOD
	INPUT	PERTURBATION			
1	WHITE NOISE UNIT COVARIANCE	$\Delta = 0$	COVARIANCE ≤ 1	$\frac{1}{2\pi} p_{22} \leq 1$	WIENER-HOPF- KALMAN (LQG)
2	$\ u\ _2 \leq 1$	$\Delta = 0$	$\ e\ _2 \leq 1$	$ p_{22} \leq 1$	L_∞/H_∞
3		$\ \Delta\ _\infty \leq 1$	BIBO STABLE	$\ p_{11}\ _\infty \leq 1$	
4		$\ \Delta\ _\infty \leq 1$, $\Delta = \text{diag}(\Delta_1, \Delta_2, \dots, \Delta_n)$	BIBO STABLE	$\ p_{11}\ _\mu \leq 1$	
5	$\ u\ _2 \leq 1$	(STRUCTURED UNCERTAINTY)	$\ e\ _2 \leq 1$	$\ p\ _\mu \leq 1$	

TABLE 3-1. ANALYSIS SUMMARY

- COMMANDS: POINTING, TRACKING, SHAPE CONTROL
- DISTURBANCES AND NOISES:
 - AERODYNAMIC
 - GRAVITY GRADIENT
 - THERMAL
 - REACTION WHEEL/CONTROL MOMENT
 - GYRO IMBALANCES AND NOISE
 - SENSOR NOISE
 - SOLAR
 - EARTH MAGNETIC
 - REACTION JET LEAKAGES
- PERTURBATIONS:
 - ACTUATOR/SENSOR UNCERTAINTY:
 - HIGH FREQUENCY DYNAMICS
 - RESIDUAL JET IMBALANCES
 - UNMODELLED STRUCTURAL DYNAMICS
 - STRUCTURAL UNCERTAINTY- MASS, STIFFNESS, DAMPING
 - STRUCTURAL CHANGES

TABLE 3-2. LSS UNCERTAINTY

Figure 3-1b. Analysis Model

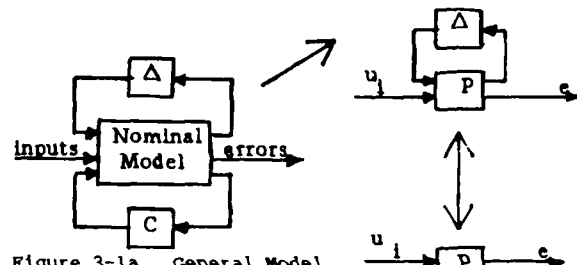


Figure 3-1a. General Model

Figure 3-1c. Synthesis Model

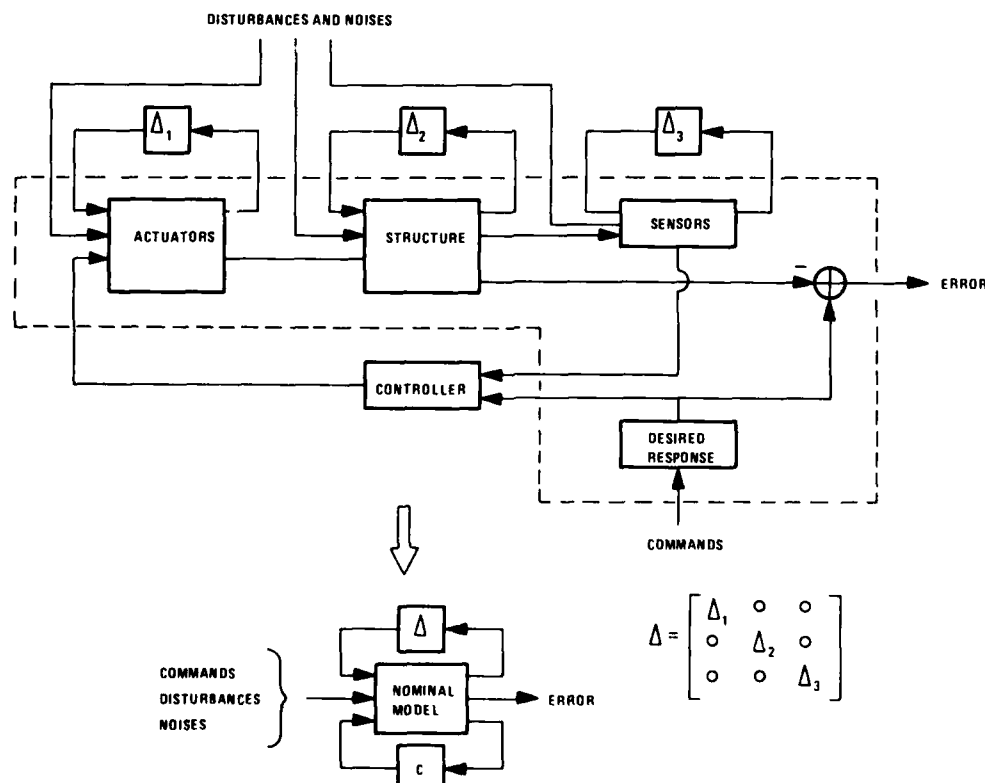


Figure 3-2. LSS Uncertainty Model

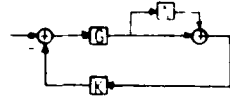
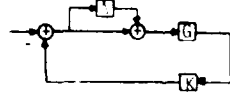
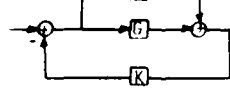
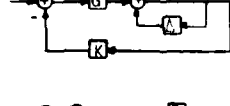
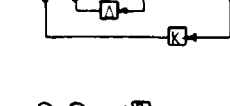
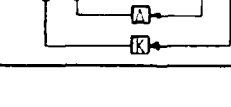
LOCATION OF Δ_1	CONDITIONS IMPOSED ON NOMINAL FEEDBACK LOOP SHAPES	REPRESENTATIVE TYPE OF UNCERTAINTY CHARACTERIZED	REPRESENTATIVE TYPE OF PERFORMANCE
	$\ [R G K (I + G K)^{-1} L^{-1}] \ < 1$	<ul style="list-style-type: none"> - output (sensor) error - neglected HF dynamics - changing numbers of rhp zero $G' = (I + \Delta)G$	<ul style="list-style-type: none"> - sensor noise effects on output response - output errors to output commands
	$\ [R K G (I + K G)^{-1} L^{-1}] \ < 1$	<ul style="list-style-type: none"> - input (actuator) errors - neglected HF dynamics - changing numbers of rhp zeros $G' = G(I + \Delta)$	<ul style="list-style-type: none"> - input response to input command
	$\ [R K (I + G K)^{-1} L^{-1}] \ < 1$	<ul style="list-style-type: none"> - additive plant errors - uncertain rhp zeros $G' = G + \Delta$	<ul style="list-style-type: none"> - input response to output command
	$\ [R (I + G K)^{-1} L^{-1}] \ < 1$	<ul style="list-style-type: none"> - LF plant parameter errors - changing numbers of rhp poles $G' = (I + \Delta)^{-1}G$	<ul style="list-style-type: none"> - output sensitivity - output errors to output commands and disturbances
	$\ [R (I + K G)^{-1} L^{-1}] \ < 1$	<ul style="list-style-type: none"> - LF plant parameter errors - changing numbers of rhp poles $G' = G(I + \Delta)$	<ul style="list-style-type: none"> - input sensitivity - input errors to input commands and disturbances
	$\ [R (I + G K)^{-1} G L^{-1}] \ < 1$	<ul style="list-style-type: none"> - LF plant parameter errors - uncertain rhp poles $G' = (G^{-1} + \Delta)^{-1}$	<ul style="list-style-type: none"> - output errors to input commands and disturbances

TABLE 4-1. REPRESENTATIVE ROBUSTNESS/PERFORMANCE CONDITIONS

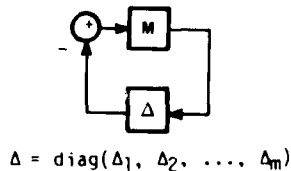


Figure 4-1. Feedback Loop as a BDBP Problem

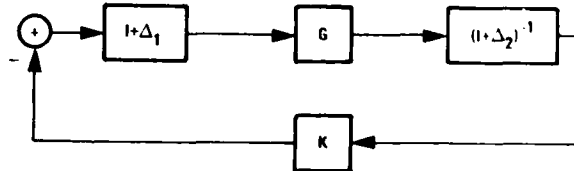


Figure 4-2. Feedback Loop with Two Uncertainties

ATTITUDE CONTROL OF LARGE FLEXIBLE SPACECRAFT

by
Götz Thieme
Dornier System
Postfach 1360
7990 Friedrichshafen 1
Germany

SUMMARY

This paper presents some results found during the investigation of control problems of large flexible spacecraft*. A triple plate configuration of such a spacecraft is defined and studied. The model is defined by modal data derived from finite element modelling. The order reduction methods applied are briefly described and results of order reduction are presented. An attitude control concept with low and high authority control has been used to design an attitude controller for the reduced model. The stability and response of the original system together with the reduced controller is analysed.

1. INTRODUCTION

For spacecraft built up to now attitude control systems and technologies have been developed, which establish conventional basics for the development of future spacecraft. Conventional spacecraft could be conceived as rigid body or as rigid body with flexible appendages. The space transportation system, together with Rendezvous & Docking, makes it possible to think of large, more complex platforms/antennas or orbital carriers. The increase in size involves a higher level of mechanical flexibility than in conventional spacecraft. For the achievement of the objectives of demanding future missions the problem of spacecraft structure and control interaction has a very high significance. The increase in size and flexibility is often combined with higher requirements for shape and attitude control. Characteristics of large flexible spacecraft with relevance for the control problem are e.g.

- Distributed parameter systems with theoretically infinite dimension and high dimension in practical approximations
- Relatively uncertain design models because of limited on-earth testing
- Many low resonant frequencies and very light, poorly known damping
- stringent accuracy requirements which require a control system bandwidth close to or above the lowest structural frequencies.

The control design for large flexible spacecraft therefore has to cope with high order systems with a variety of sensors and actuators, uncertain control plant parameters and hardware limitations, such as control device design and small on board computers. The attitude control designer also has to consider the extreme relation to structural mechanics. This paper concentrates on theoretical aspects of the attitude control problem of large flexible spacecraft. It reports on baseline investigations for the test and comparison of design approaches and techniques.

2. STUDY APPROACH

The objectives of the study carried out are the investigation of control problems of large flexible spacecraft and the identification of basic problems that will be studied experimentally. The first steps have been the analysis and definition of requirements for the attitude control design and the dynamics model set up. Then order reduction methods have been applied to test the reduction methods and to achieve reduced order models. The next step is the controller design itself, but it is clear that order reduction and controller design have to be seen as elements of an iteration loop, in order to come to a final reduced order controller. The design is finished by a performance evaluation and critical assessment of the designed systems and the design methods applied. The identification of basic problems leads to the definition of an appropriate test set up and test program to study those problems experimentally. Fig. 2-1 shows the study task flow in phase 1. The study is still under work. This paper presents results found up to the major part of the controller design task.

*This work has in main parts been done under ESA contract 5310

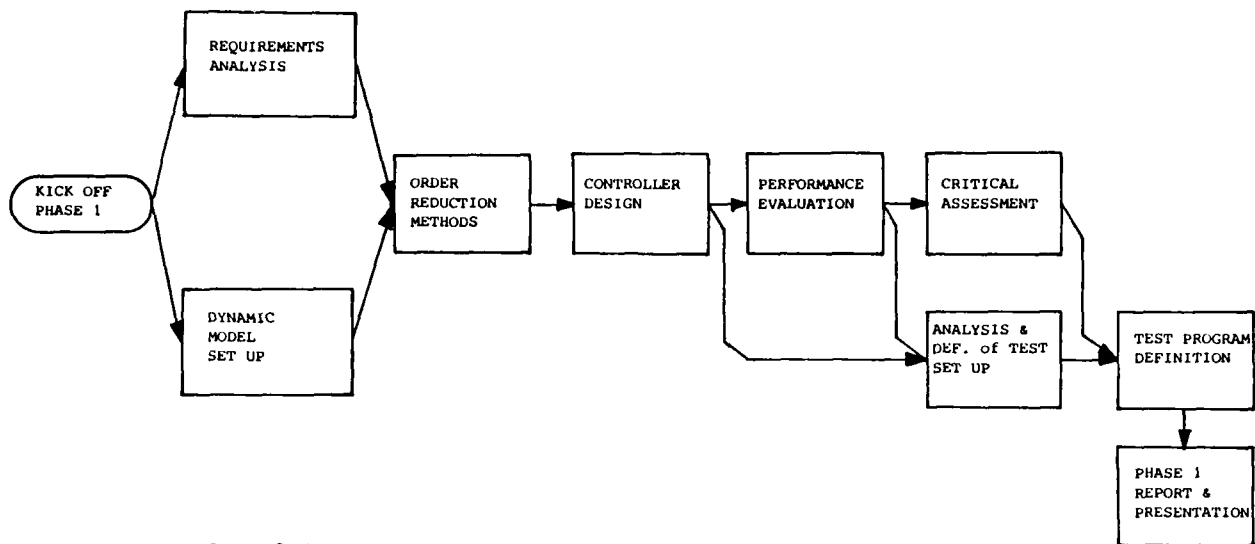


FIG. 2-1: STUDY TASK FLOW PHASE 1

3. DISTRIBUTED FLEXIBILITY MODEL

Three models were used as reference spacecraft in the studies. The models do not represent real or planned spacecraft, but they are representative for the characteristics which play a role in the design of attitude control systems for future large spacecraft. The simplest model is a free beam. The second model has been proposed in the literature as a generic model. It is a flat plate with a rigid centre body and has been developed at Purdue University. The third model is constructed of three coupled plates with the goal to have a spacecraft with nonhomogeneous flexibility. This three plate model proved to be the most interesting model, giving the most design problems. We will therefore concentrate on this model. The model consists of a central plate, having higher stiffness and mass density, which is coupled to two identical plates with lower eigenfrequencies. The modal data of this triple plate configuration has been derived from a finite element model. The physical size is 100 m by 150 m. Each of the three plates has a width of 50 m.

The thickness of the inner plate is set to 2 m, whereas the outer plates are 1 m thick. The elasticity modulus is $2 \cdot 10^{11}$ N/m² for the outer plates and $7 \cdot 10^{11}$ N/m² for the centre plate. The mass density is 2 kg/m³. These values are quite arbitrary. They have been selected in order to get densely packed frequencies starting at about 0.02 Hz. From the 40 element model with 171 degrees of freedom the following 20 lowest eigenfrequencies were found:

Mode	Omega [rad/s]	Frequency [Hz]
1	0	0
2	0	0
3	0	0
4	0.139	0.022
5	0.146	0.023
6	0.269	0.043
7	0.281	0.045
8	0.292	0.047
9	0.311	0.050
10	0.483	0.077
11	0.628	0.100
12	0.668	0.106
13	0.763	0.121
14	0.793	0.126
15	0.972	0.155
16	1.001	0.159
17	1.173	0.187
18	1.243	0.198
19	1.259	0.200
20	1.459	0.232

TABLE 3-1: EIGENFREQUENCIES OF THE DISTRIBUTED FLEXIBILITY MODEL

Plots of those nine modes which were identified as dominant are presented in Fig. 3-1 to 3-3. Input/Output transfer functions of the model are evaluated directly from the eigenvalue and eigenvector solutions of the free-free structure as shown below in (1) to (3), see ref. 1.

$$y(k) = \sum_j T_u(k,j)u(j) + \sum_l T_{\text{dist}}(k,l)\text{dist}(l) \quad (1)$$

with $j = 1, \dots, \text{NB}$
 $l = 1, \dots, \text{NDIST}$
 $k = 1, \dots, \text{NC}$

$$T_u(k,j) \triangleq \sum_i \frac{CD_i(k)s + CP_i(k)}{s^2 + 2\zeta_i\omega_i s + \omega_i^2} B_i(j) \quad (2)$$

$$T_{\text{dist}}(k,l) \triangleq \sum_i \frac{CD_i(k)s + CP_i(k)}{s^2 + 2\zeta_i\omega_i s + \omega_i^2} \text{DIST}_i(l) \quad (3)$$

with $i = 1, \dots, \text{NMR}$

where y : measurement vector

u : control vector

dist : disturbance vector

NMR : number of retained modes

ζ_i : modal damping

f_i : modal frequency ($\omega_i = 2\pi f_i$)

$CP_i(k)$ is defined by:

$$CP_i(k) \triangleq \frac{ds_i(k)}{M_i}; \quad k=1, \dots, \text{NC}, i=1, \dots, \text{NMR} \quad (4)$$

$ds_i(k)$ is the i 'th mode's displacement at sensor station k ; M_i is the i 'th mode's generalized mass.

$B_i(j)$ and $\text{DIST}_i(l)$ are defined by

$$B_i(j) \triangleq da_i(j) \quad j=1, \dots, \text{NB}; \quad i=1, \dots, \text{NMR}; \quad (5)$$

$$\text{DIST}_i(l) \triangleq da_i(l) \quad l=1, \dots, \text{NDIST}; \quad i=1, \dots, \text{NMR}; \quad (6)$$

where $da_i(j)$, $da_i(l)$ are the i 'th mode's displacements at actuator station j and disturbance station l respectively.

The $CD_i(k)$ terms in (2) and (3) are used for rate type sensors and are defined as in (4) if rate sensor signals are to be modelled.

Equation (2) and (3) may be written in modal form by defining proportional and derivative modal gains:

$$CPU_i(k,j) \triangleq CP_i(k) \cdot B_i(j) \quad (7)$$

$$CDU_i(k,j) \triangleq CD_i(k) \cdot B_i(j) \quad (8)$$

$$CPDIST_i(k,l) \triangleq CP_i(k) \cdot \text{DIST}_i(l) \quad (9)$$

$$CDDIST_i(k,l) \triangleq CD_i(k) \cdot \text{DIST}_i(l) \quad (10)$$

giving:

$$T_u(k,j) = \sum_i \frac{CDU_i(k,j)s + CPU_i(k,j)}{s^2 + 2\zeta_i\omega_i s + \omega_i^2} \quad (11)$$

$$T_{\text{dist}}(k,l) = \sum_i \frac{CDDIST_i(k,l)s + CPDIST_i(k,l)}{s^2 + 2\zeta_i\omega_i s + \omega_i^2} \quad (12)$$

The transfer function (1) with T_u and T_{dist} expressed as in (11) and (12) consists of parallel second order terms with resonant frequencies according to the eigenfrequencies and weight factors according to the modal gains. The transfer function can be expressed in state space form by decoupled second order state space forms for each mode. Hence the system matrix consists of a band of 2 by 2 matrices along their diagonal.

So far each node of the structural model can be defined as input or output station. For the use of the model in the studies different sensor/actuator stations have been defined. The following configuration proved to be most reasonable:

- . Rotation sensor X, y at the centre node
- . Translation sensors z and force actuators z at the edges of the outer plates

The centre node is considered as a reference for attitude control.

4. ORDER REDUCTION

The idea of order reduction is to try to approximate a system of high order by a lower order system in order to ease the simulation, analysis and controller design. The design based on a reduced order system may also be less costly and reduces the controller order, which may be important for the implementation in the control computer aboard a spacecraft. The methods for order reduction of state representations may be classified into four groups:

- minimization of output error
- minimization of equation error
- modal reduction
- singular perturbation.

The finite element model itself is already a reduced model and the original system order is determined by the number of eigensolutions computed from the FEM. In our case the original order is 40. The mathematical methods of order reduction are then applied to this 40th order system. Since our model is based on modal data it appears reasonable to use modal methods for order reduction.

4.1 Modal Gain Considerations

In section 3 the transfer function of the model has been expressed in form of parallel second order oscillators. Each oscillator represents the contribution of a mode to the input/output transfer function. The weight is represented by the modal gains (7) to (10). From this consideration it is possible to select those modes which contribute most to the transfer function as dominant and keep them in a reduced order model. Modes which contribute little are neglected. This approach has been used for a reduced model of the Spacelab Instrument Pointing System / ref. 1 /.

In simulation models for disturbance response analysis the modal gains in (11) and (12) should be considered, but for control studies only the modal gains from control input to output are relevant.

4.2 Litz' Modal Approach

Modal approaches to order reduction require a transformation of the original system to the Jordan canonical form and take only the dominant eigenvalues of the original model into a reduced model. Methods have been developed since the mid sixties. The techniques of E.J. Davison, S.A. Marshall and M.R. Chidambara in their original versions have some shortcomings, which have been avoided by a newer technique of L. Litz (ref. 2). In reference 3 a unified derivation and critical review of modal approaches has been given. Litz' technique is shown to be both highly accurate and powerful. Therefore this technique has been selected for application to the flexible spacecraft model. Litz' method links modal techniques to least square reduction techniques. It gives an explicit formulation for the reduced model and is optimal in the sense that it minimizes the weighted integral of the square of the deviations between the original and the reduced model.

The original states are sorted such as to have the relevant states in a subvector x_1 :

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + B u \quad (13)$$

The relevant states may be identified by their physical meaning or it may be the states which contribute to the output vector:

$$y = [C_1 \ 0] x_1 \quad (14)$$

The reduced model is an approximation of the substates x_1 :

$$\dot{\tilde{x}}_1 = A_R \tilde{x}_1 + B_R u. \quad (15)$$

As for all modal methods the original system (13) is transformed by:

$$x = V \cdot z = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} z \quad (16)$$

to Jordan canonical form

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} \mathcal{A}_1 & 0 \\ 0 & \mathcal{A}_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} B_1^* \\ B_2^* \end{bmatrix} u \quad (17)$$

where

$$\begin{bmatrix} B_1^* \\ B_2^* \end{bmatrix} = B^* = V^{-1} B.$$

Now again the transformed state vector is reordered such that z_1 represents the dominant part, found by certain dominance measures. The structure of the original model is shown in Fig. 4-1.

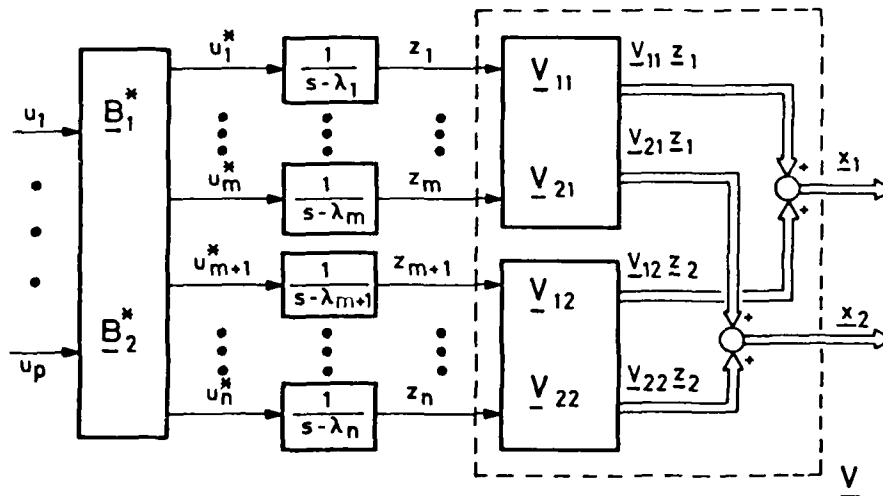


FIG. 4-1: STRUCTURE OF ORIGINAL MODEL

The reduced model is not derived by neglecting the nondominant eigenmotions z_2 , but by reconstructing the z_2 out of the retained states z_1 (see Fig. 4-2):

$$\tilde{z}_2 = E z_1 \quad (18)$$

$$\text{giving } \tilde{x}_1 = (V_{11} + V_{12} E) z_1 = M z_1 \quad (19)$$

and the reduced model (15) as

$$\dot{\tilde{x}}_1 = M \mathcal{A}_1 M^{-1} \tilde{x}_1 + M B_1^* u. \quad (20)$$

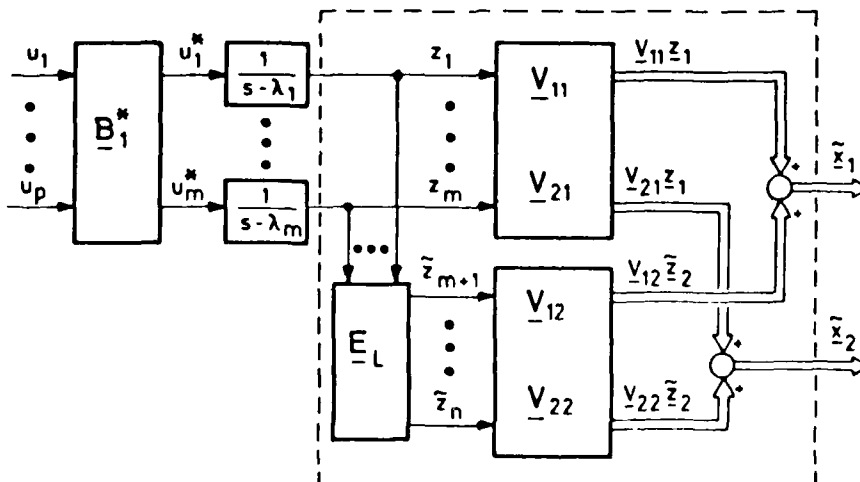


FIG. 4-2: REDUCED MODEL OF LITZ

The matrix E is determined such that the following performance index is minimized:

$$I = \sum_{i=1}^r q_i^2 \int_0^{\infty} \epsilon_i(t)^2 dt$$

with $\epsilon_i(t) = z_{2i}(t) - \tilde{z}_{2i}(t)$

r number of inputs

The index i indicates that only the i -th input is excited by a unity step. The weight factors q_i allow different weighting of the system inputs.

The optimization problem can be solved explicitly for

$$E = E(B^*, \Lambda, q_1, \dots, q_r).$$

The solution also guarantees stationary accuracy by

$$\epsilon_i(t \rightarrow \infty) = 0 \text{ for } i = 1, r.$$

Litz has also proposed a method of generating a dominance measure. This measure does not only take into account the stability of the eigenvalue, i.e. its distance to the imaginary axis, but in addition its controllability and observability characteristics. To keep the presentation clear a single input single output system with single eigenvalues $\lambda_1, \dots, \lambda_n$ is considered. The state space description in Jordan canonical form is

$$\begin{aligned} \dot{z} &= \Lambda z + b^* u \\ y &= c^* \cdot z \end{aligned}$$

The vector elements b_i^* and c_i^* determine the controllability and observability of the system.

For $b_i^* = 0$ the eigenmotion z_i is not controllable with u , and for $c_i^* = 0$ it is not observable from y . The unity step response can be expressed as

$$y = \sum_{k=1}^n \frac{b_k^* c_k^*}{\lambda_k} (e^{\lambda_k t} - 1), \quad t > 0.$$

From this

$$q_k = \frac{|b_k^* c_k^*|}{|\lambda_k|}$$

can be taken as a measure of the influence of eigenvalue λ_k on the dynamic behavior of the system. The definition of the equivalent measure for multiple input multiple output systems is easy, considering each signal path from input u_i to output y_j :

$$q_{ijk} = \frac{|b_{kj}^* c_{ik}^*|}{|\lambda_k|}$$

If this q_{ijk} is related to the stationary value $y_{ij\infty}$, the dimensionless value

$$r_{ikj} = \frac{|q_{ikj}|}{|y_{ij\infty}|}$$

gives a measure how the response $y_{ij}(t)$ is influenced by the eigenvalue λ_k . As a first criterion it can be stated that eigenvalues with small r_{ijk} are not dominant within the path (i,j) . However also eigenvalues with high r_{ijk} can be not dominant, if dynamic compensation occurs. The compensation can be seen from small values of the amplitude $/G_{ij}(j/\lambda_k)/$. Therefore the measure r_{ijk} is multiplied by

$$A_{ij}(\lambda_k) = \frac{|G_{ij}(j/\lambda_k)|}{|G_{ij}(0)|}$$

The weighted dominance measure becomes

$$\hat{r}_{ijk} = r_{ijk} A_{ij}(\lambda_k).$$

A dominance measure, which also reflects that the output is influenced differently from the inputs, which may also have different amplitudes, can be defined either by

$$d_{ijk} = r_{ijk} \cdot N_{ij}$$

$$\text{or } \hat{d}_{ijk} = \hat{r}_{ijk} \cdot N_{ij}$$

with

$$N_{ij} = \frac{y_{ij} u_{imax}}{\max_{i=1,r} (y_{ij} u_{imax})}$$

The evaluation of the dominance with respect to the overall system can be done by either the maximum or the sum of all measures d_{ijk} which belong to the eigenvalue λ_k .

$$M_k = \max_{(i,j)} d_{ijk}$$

$$\hat{M}_k = \max_{(i,j)} \hat{d}_{ijk}$$

$$S_k = \sum_{i,j} d_{ijk}$$

$$\hat{S}_k = \sum_{i,j} \hat{d}_{ijk}$$

In reference 2 Litz formulates two criteria:

- 1) The eigenvalues λ_k with the highest measures M_k , S_k are dominant with respect to controllability / observability.
- 2) The eigenvalues λ_k with the highest measures \hat{M}_k , \hat{S}_k are dominant with respect to the transfer behavior.

For reduced models to be utilized for the design of reduced controllers the eigenvalues should be selected according to criterion 1.

4.3 Results of Order Reduction

a) Determination of Dominant Eigenvalues

From the modal gains the following dominance classification has been derived:

Rank	Flex. Mode No.	ω_k [rad/s]	Modal Gain/ ω_k^2
1	1	0.139	1.8865 E-2
2	2	0.146	8.5103 E-3
3	3	0.269	5.3749 E-3
4	6	0.311	5.3658 E-3
5	4	0.281	3.2024 E-3
6	5	0.292	3.1510 E-3
7	7	0.483	2.8903 E-3
8	10	0.763	1.4389 E-3
9	9	0.668	1.4123 E-3
10	11	0.793	1.2404 E-3
11	13	0.972	9.9164 E-4
12	15	1.243	8.9368 E-4
13	16	1.259	7.5801 E-4
14	8	0.628	6.2639 E-4
15	12	0.972	5.7491 E-4
16	14	1.173	3.75583E-4
17	17	1.459	1.4874 E-4

TABLE 4-1: DOMINANCY ORDER DERIVED BY MODAL GAIN CONSIDERATION

The question now is how much modes shall be retained. From Table 4-1 it can be seen that the measure decays rapidly from the highest value down to about the 8th or 9th. If more modes would be taken into the reduced model then it becomes more difficult to separate dominant from nondominant modes, because the measure is very close for all modes. Therefore it seems reasonable to retain the first 9 Modes of Table 4-1.

The results of the dominance analysis using Litz' technique are presented in Table 4-2. It gives the dominance measures M_k , S_k with respect to controllability/observability and \hat{M}_k , \hat{S}_k with respect to the transfer behavior. As for the modal gain the measures decay rapidly for the most dominant nine modes. This seems to be a rational to take nine elastic modes into the reduced model. The ranking of modes is different for all dominance measures, but the nine dominant modes are the same for the modal gain method, for S_k and \hat{S}_k . Table 4-3 compares the ranking of the retained modes. Reduced models with less than 9 modes were not satisfying in response analysis and during the design of active vibration damping.

No.	Eigenvalues		Contr./Obser. Dominancy		I/O Transfer Dominancy	
	REAL	IMAG	MAXIMUM	SUM	MAXIMUM	SUM
1	-0.0007	0.1386	0.0163	0.5193	4.3135	72.9620
2	-0.0007	-0.1386	0.0163	0.5193	4.3135	72.9620
3	-0.0007	0.1462	0.0078	0.3539	3.3939	43.1744
4	-0.0007	-0.1462	0.0078	0.3539	3.3939	43.1744
5	-0.0013	0.2688	0.0431	0.6362	8.0551	100.0000
6	-0.0013	-0.2688	0.0431	0.6362	8.0551	100.0000
7	-0.0014	0.2811	0.0290	0.2554	10.5785	75.4626
8	-0.0014	-0.2811	0.0290	0.2554	10.5785	75.4626
9	-0.0015	0.2922	0.0044	0.1932	2.4384	20.2673
10	-0.0015	-0.2922	0.0044	0.1932	2.4384	20.2673
11	-0.0016	0.3110	0.0531	0.3124	7.0587	31.8490
12	-0.0016	-0.3110	0.0531	0.3124	7.0587	31.8490
13	-0.0024	0.4830	0.0058	0.1509	8.7702	51.9795
14	-0.0024	-0.4830	0.0058	0.1509	8.7702	51.9795
15	-0.0031	0.6279	0.0054	0.0691	0.4246	3.1283
16	-0.0031	-0.6279	0.0054	0.0691	0.4246	3.1283
17	-0.0033	0.6685	0.0028	0.1223	2.0986	13.9860
18	-0.0033	-0.6685	0.0028	0.1223	2.0986	13.9860
19	-0.0038	0.7630	0.0109	0.1713	2.1727	15.8973
20	-0.0038	-0.7630	0.0109	0.1713	2.1727	15.8973
21	-0.0040	0.7934	0.0011	0.0317	0.0189	0.3278
22	-0.0040	-0.7934	0.0011	0.0317	0.0189	0.3278
23	-0.0049	0.9716	0.0011	0.0206	0.3547	2.0379
24	-0.0049	-0.9716	0.0011	0.0206	0.3547	2.0379
25	-0.0050	1.0006	0.0057	0.0886	0.1431	1.7979
26	-0.0050	-1.0006	0.0057	0.0886	0.1431	1.7979
27	-0.0059	1.1733	0.0042	0.0452	0.1497	1.2123
28	-0.0059	-1.1733	0.0042	0.0452	0.1497	1.2123
29	-0.0062	1.2432	0.0018	0.0291	0.7894	4.7212
30	-0.0062	-1.2432	0.0018	0.0291	0.7894	4.7212
31	-0.0063	1.2589	0.0015	0.0377	0.5670	3.4694
32	-0.0063	-1.2589	0.0015	0.0377	0.5670	3.4694
33	-0.0073	1.4587	0.0002	0.0088	0.0049	0.0412
34	-0.0073	-1.4587	0.0002	0.0088	0.0049	0.0412

Normalization: 1251.78

TABLE 4-2: RESULTS OF DOMINANCY ANALYSIS BY LITZ' METHOD

Rank	Dominancy Measure		
	\hat{S}_k	S_k	Modal Gain
1	3	3	1
2	4	1	2
3	1	2	3
4	7	6	6
5	2	4	4
6	6	5	5
7	5	10	7
8	10	7	10
9	9	9	9

TABLE 4-3: RANKING OF ELASTIC MODES BY THREE DOMINANCY MEASURES

b) Reduced Models

The reduced model derived from the modal gain method is achieved by simply neglecting the terms for the neglected modes in the transfer function (1). The same reduced model, caused by the model structure, is found applying Litz' technique without optimization. The approximation of the response to unity step forces at all force actuators of the original model is shown in Fig. 4-3 and 4-4. Only elastic modes are contained. The final reduced model has to be augmented by the rigid body modes, which have to be retained in any case and which are therefore not considered in the dominancy analysis.

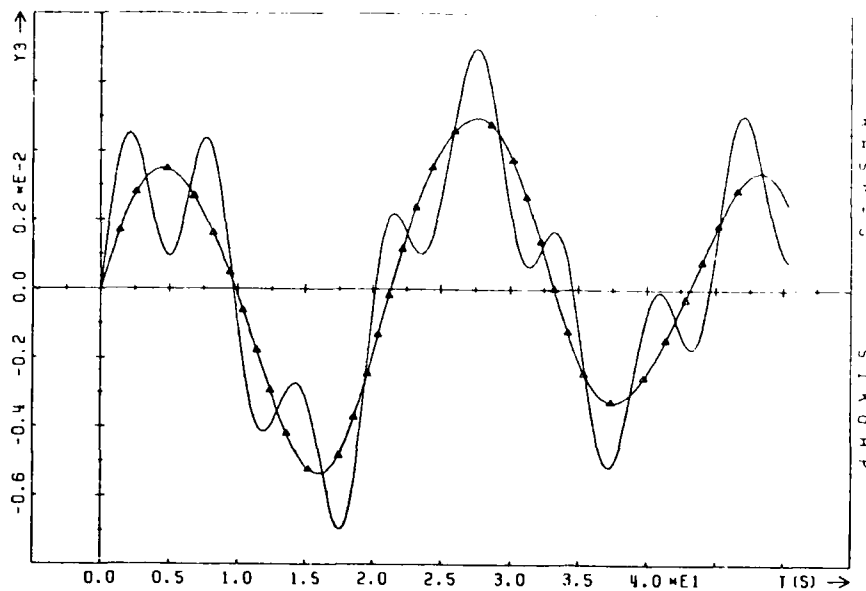


FIG. 4-3: RATE AT INNER EDGE OF OUTER PLATE

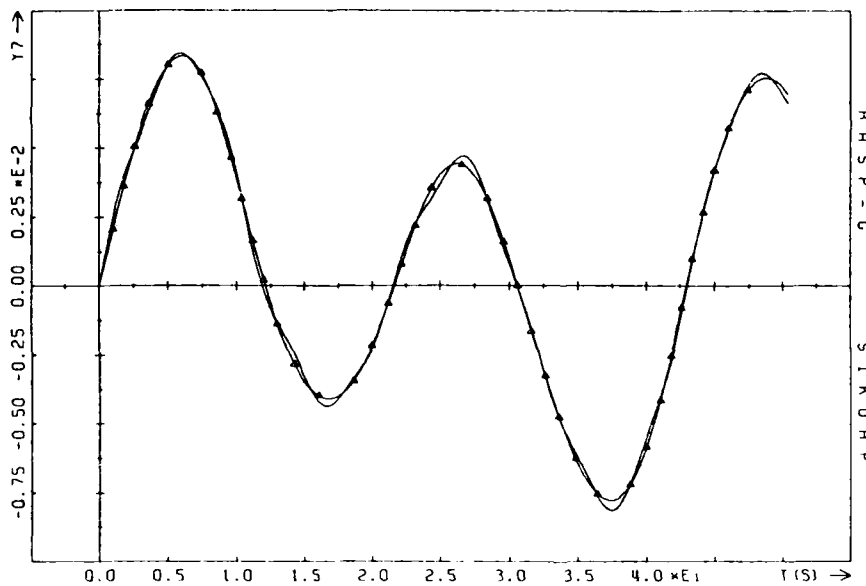


FIG. 4-4: RATE AT OUTER EDGE

Unfortunately the generation of an optimally reduced model by Litz' technique was not successful. In order to achieve the presumed system description (13) and (14) the original model has been transformed such that the output vector becomes part the state vector, i.e. C_1 in (14) becomes a unity matrix. The problem arose then from the transformed original system. First the resulting Jordan canonical form was too inaccurate for numerical reasons, such that the reduced models derived from the canonical form had wrong, partially unstable, eigenvalues. This could be solved by application of a better transformation scheme, however the computation of the optimal matrix E for reconstruction of the neglected state variables failed.

The dominance measures of Litz proved to be a good tool for the selection of eigenvalues to be retained, but the flexible spacecraft model could not be reduced optimally in the sense of Litz. Since the algorithm has given good results in many cases, the flexible spacecraft model seems to be a really critical system for Litz' technique.

5. REQUIREMENTS FOR CONTROLLER DESIGN

In a theoretical study on control design techniques it is difficult to define requirements which are design drivers. In practical design cases mainly the performance requirements disturbances and realization constraints determine the basis on which a control system is designed and verified. An accuracy requirement can often also be formulated as a control system bandwidth requirement. In the context of elastic modes of the plant model the design bandwidth is an important factor. Problems can be expected if the control bandwidth is close to the structural eigenfrequencies, or if even eigenfrequencies of the plant fall within the control bandwidth. To analyse these problems a control system bandwidth in the range of the lowest structural eigenfrequencies has been set as a design requirement for the attitude control. The controller complexity is another critical area with respect to implementation on flight computers, number of sensors/actuators and so on. As a general requirement the controller complexity should be kept to a minimum and a controller implementation with microprocessor based electronics should be possible.

6. CONTROLLER DESIGN

Based on the experience that control problems with elastic modes are largely reduced with increased damping and based on results in other studies, the controller is separated in two tasks:

- Low Authority Control (LAC) for vibration damping
- High Authority Control (HAC) for attitude control.

The LAC is designed to increase modal damping by velocity feedback. As a design goal the modal damping should be at least 5% due to the LAC. The LAC design has been done by Dr. G. Schulz at DFVLR. It is based on a dissipation energy minimization approach. The LAC gains L for local feedback of the velocities at the edges of the outer plates have been derived using the original model. The design goal of 5% damping of each mode has been reached.

The HAC design is done for the reduced model with LAC:

$$\begin{array}{ll} \text{Reduced model} & \dot{x}_N = A_N x_N + B_N u \\ \text{Velocity measurements} & y_L = C_{LN} x_N \\ \text{LAC} & u_L = -L y_L \end{array}$$

The control plant for the HAC design then becomes

$$\begin{aligned} \dot{x}_N &= (A_N - B_N L C_{LN}) x_N + B_N u_H \\ y &= C_N x_N \end{aligned}$$

The controller for this system consists of a state estimator and a linear feedback control law

$$u_H = -k \hat{x}_N.$$

The state estimator has the form

$$\begin{aligned} \dot{\hat{x}}_N &= (A_N - B_N L C_{LN} - G C_N - B_N k) \hat{x}_N + G y \\ \hat{x}_N(0) &= 0. \end{aligned}$$

The estimator gains G are determined as stationary Kalman filter gains. The control gains k are determined as if the state x_N would be present instead of \hat{x}_N . The steady-state regulator approach is used with a quadratic performance index

$$J = \int_0^{\infty} (x_N^T Q x_N + u_H^T R u_H) dt.$$

The weighting matrix R is set to unity and the matrix Q is set as a diagonal matrix. The elements of Q are chosen such as to achieve closed loop eigenvalues of the rigid body rotation above the first two flexible modes. No increase in damping of the flexible modes by HAC has been required. The setting of the Q matrix to achieve the desired eigenvalues is easily done with few iterations because the relation between eigenvalues and states is known from the special model structure.

The observability and controllability conditions have to be observed in the design. Finally the HAC and observer for the reduced system have to be coupled to the original model in order to evaluate stability when spillover from neglected modes is present.

In a first step the HAC has been designed without LAC. The original system is stable, i.e. the neglected modes do not affect stability. This result confirms the proper choice of dominant modes to be retained in the reduced model. The elastic modes however are still lightly damped. Fig. 6-1 shows the rotation of the centre about x after 1 degree initial attitude deviation. The slow oscillation is the bending mode 5. Fig. 6-2 shows the rotation of the centre about y after 0.65 degrees initial attitude deviation.

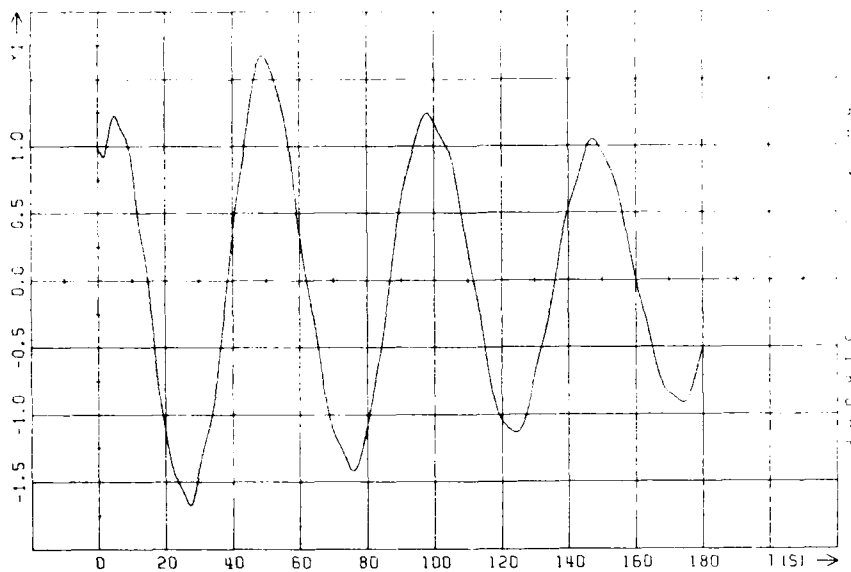


FIG. 6-1: ϕ_x RESPONSE OF THE CENTRE WITH HAC ONLY

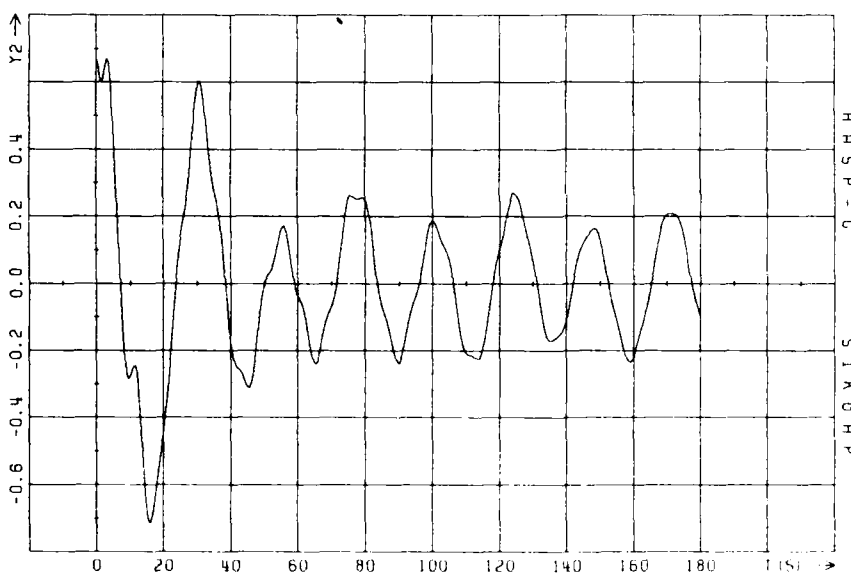
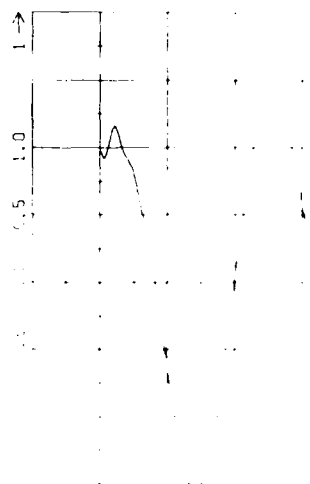
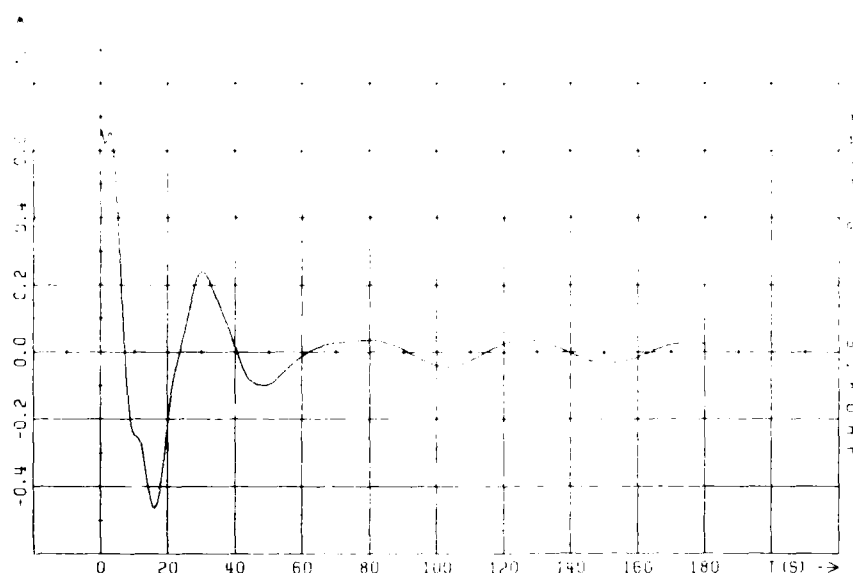


FIG. 6-2: ϕ_y RESPONSE OF THE CENTRE WITH HAC ONLY

The next step was to include LAC into the HAC design as described above. The resulting HAC gains are slightly lowered. The damping of higher modes is increased, but the slow modes get only slightly higher damping. The initial increase in damping by LAC is decreased by the attitude controller. The actuator amplitudes are reduced and the responses become smoother. Fig. 6-3 shows the response like Fig. 6-1 but now with LAC. The bending mode 5 still dominates. Fig. 6-4 presents the ϕ_y response as Fig. 6-2 but now with LAC. In this response the smoothing effect is obvious. The higher modes are well damped, only the slow mode 4 remains visible.

FIG. 6-3: ϕ_x RESPONSE OF THE CENTRE WITH HAC AND LACFIG. 6-4: ϕ_y RESPONSE OF THE CENTRE WITH HAC AND LAC

7. CONCLUSION

The triple plate spacecraft model proved to be a challenging reference plant for the investigation of order reduction and controller design. The order reduction approach which retains modes with high gains in the input/output transfer function resulted in a satisfactory reduced model. Its application is straight-forward and integrated into the process of generating state space forms out of the modal data from finite element models. The modal reduction technique of Litz, which has been excellent in other applications, fails because of numerical problems with the flexible spacecraft model. The dominance measures of Litz however are well suited for the selection of eigenvalues to be retained in a reduced model.

The control concept with low authority controller for vibration damping and high authority controller for attitude control proved to be appropriate. For the triple plate model it was possible to design also an attitude controller without additional vibration damping. But the increase in modal damping in general eases the attitude controller design, reduces the sensitivity to model uncertainties, and smoothes the system response and controller action.

Since the reduced order controller was not affected by spillover, a further order reduction of the model with added vibration damping seems possible.

References

- 1) Thieme G., P. Rogers, D. Sciacovelli
Sampled Control Stability of the ESA Instrument Pointing System, IFAC/ESA Symposium on Automatic Control in Space, Noordwijkerhout, The Netherlands, 1982.
- 2) Litz L.
Reduktion der Ordnung linearer Zustandsraummodelle mittels modaler Verfahren, Hochschulverlag Stuttgart, 1979.
- 3) Bonvin D., D.A. Mellichamp
A unified Derivation and Critical Review of modal approaches to Model Reduction, Int. J. Control, 1982, Vol. 35, No. 5, 829-848.

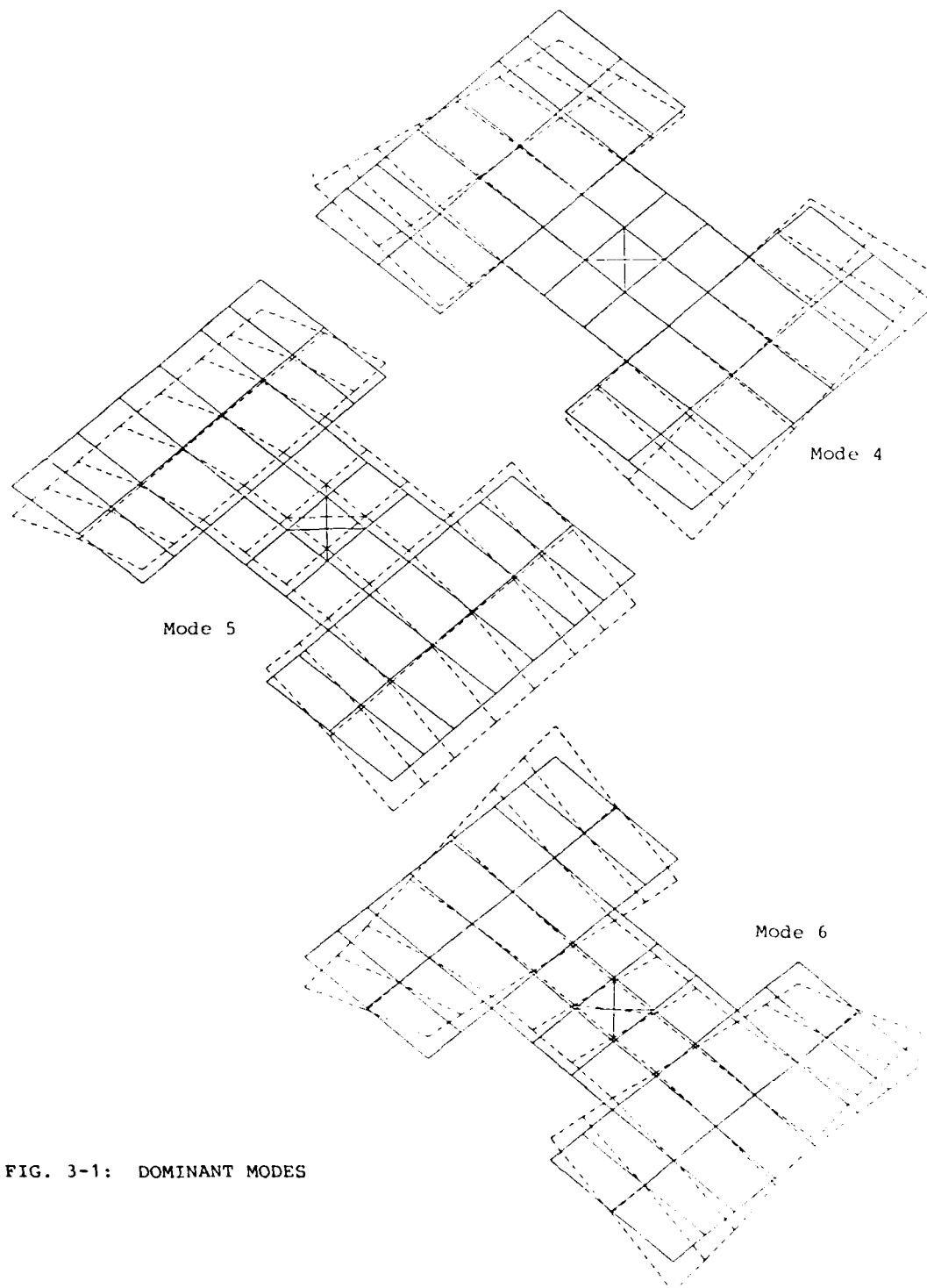


FIG. 3-1: DOMINANT MODES

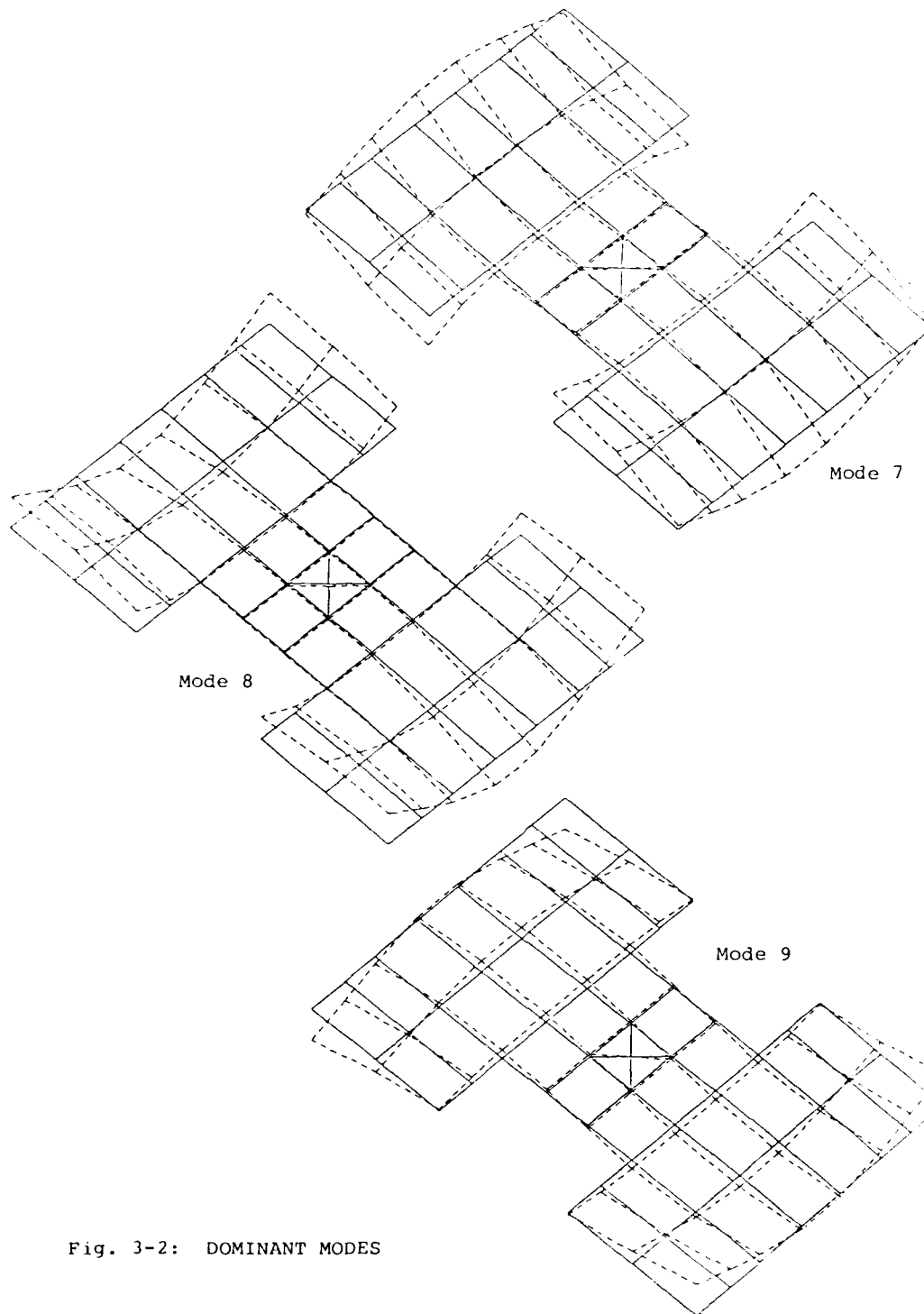


Fig. 3-2: DOMINANT MODES

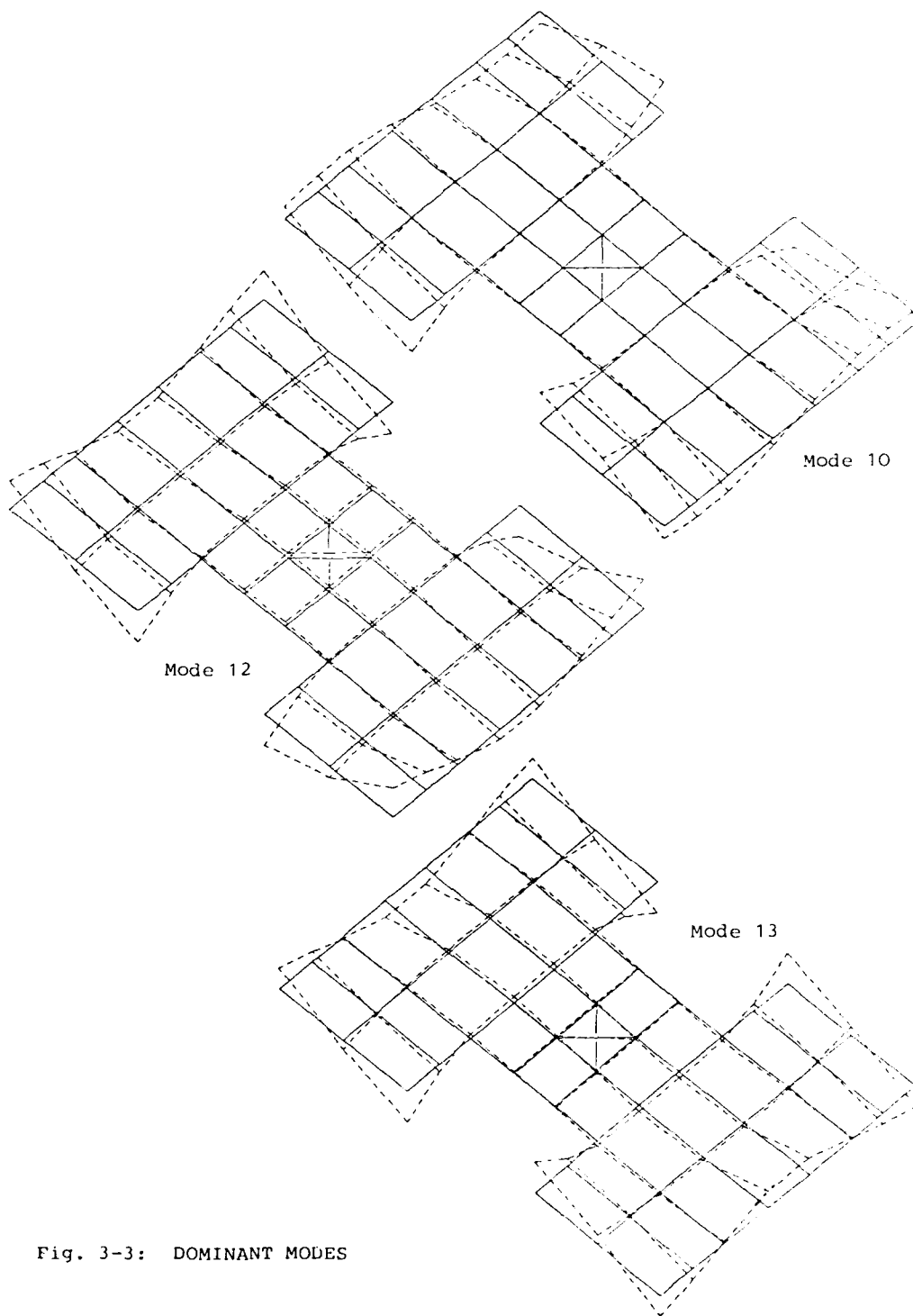


Fig. 3-3: DOMINANT MODES

LOW AUTHORITY CONTROL OF FLEXIBLE SPACECRAFT VIA NUMERICAL OPTIMIZATION

Gerd Schulz

DFVLR-Institut für Dynamik der Flugsysteme

Oberpfaffenhofen

D-8031 Wessling

F. R. Germany

Summary

For the design of low authority controllers for attitude control of flexible spacecraft a new method based on numerical optimization is presented. Thereby a dominant criterion to be optimized is determined, and the design requirements are formulated as constraints. Two different criteria are investigated: first a dissipation energy criterion is maximized and second a control energy criterion is minimized. The method is applied for attitude controller design of two models, the "Draper Model No. 1" and the "Purdue Model". The first represents a flexible tetrahedral truss structure while the latter can be considered as model of a solar power satellite. Damping requirements for these models are formulated as design objectives and the controllers are determined using the proposed method. The achieved results show that due to the now existing efficient optimization software, controller design via numerical optimization is an useful and flexible tool for system design. This flexibility arises from the fact that arbitrary design criteria can be implemented.

1. Introduction

Future space missions like large communication platforms, orbital carriers, advanced space processing facilities and solar power systems require large light weight structures. Size and flexibility of these space structures will require new shape and attitude control systems.

These large flexible spacecrafts are mainly two or three - dimensional truss structures. For the determination of the dynamic behaviour of these structures finite element methods have to be employed, to derive the eigenfrequencies and eigenmodes. A great number of these structural modes are falling within the control system bandwidth. Therefore specific controllers have to be developed to fulfil the attitude control requirements.

One controller design methodology is to increase the damping of the eigenmotion via output feedback (low authority control) in order to ease the design of the specific attitude controller (high authority control). The low authority controller design is performed for reduced order models of moderate dimension and shall guarantee an improved damping of the complete model. The high authority controller design which may employ special mode cancellation techniques is based on reduced order models of small dimension to fulfil the attitude control requirements.

Here the low authority controller design part is investigated in more detail. This low authority control design was first proposed by Aubrun [1] using perturbation techniques. This method efficiently allows to improve the damping of the eigenmodes. If additional design goals as e.g. positioning of actuators and sensors, minimum gain, lower/upper bounds of damping have to be fulfilled, numerical optimization can be applied successfully [2].

A low authority controller design is presented which is based on nonlinearly constrained numerical optimization. Thereby the design requirements are formulated as constraints of the optimization. A dominant optimization criterion has to be determined, which supports the optimization routine to fulfil the design requirements. Two different criteria are investigated in the following: first a dissipation energy criterion is maximized and second a control energy criterion is minimized.

The investigations are performed for two typical models of flexible space structures. The first is a flexible truss structure known in the literature as "Draper Model No. 1" [3] and the second is a flexible plate in orbit known as "Purdue Model" [4].

2. Maximization of Dissipation Energy

Low authority controller design via maximization of the dissipation energy is demonstrated for the "Draper Model No. 1", presented in the following:

2.1 "Draper Model No. 1"

This model is a tetrahedral truss structure. Its design goal was to retain many of the characteristics of a typical large space structure, and at the same time, keep the order of the problem small (less than 20 modes). The resulting model shown in Fig. 2.1 meets these design requirements.

AD P003 396

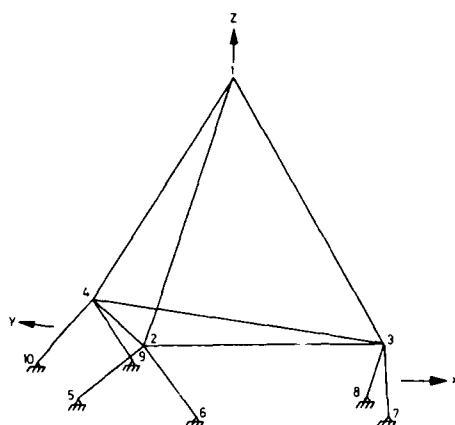


Figure 2.1 Tetrahedral finite-element model

The structure is similar to typical optical or radar systems. The truss structure is supported by three right-angled bipods. The bipod legs are pinned to ground, and all other nodes are clamped. The individual truss members, including the bipods, have elastic flexibility in their axial direction only, i.e. they can undergo compression and elongation but no lateral bending. The four vertices of the tetrahedron each have three degrees of freedom, so that the entire model can be completely represented by 12 structural modes, which are shown in appendix C.

The structural model is described in consistent, but unspecified units. The edges of the (regular) tetrahedron are each 10 units long, and the bipod legs are each $2\sqrt{2}$ units long. (The "horizontal" base of the tetrahedron is therefore situated 2 units above ground, i.e., above the x,y plane.) The top vertex is defined as LOS (line-of-sight), and the six bipod legs are defined to act as "member dampers," i.e., active axial spring/dashpots, each equipped with a rate and position sensor. The three bipods represent therefore 6 colocated actuator/sensor pairs, each pair being identified with a bipod leg. Sensor measurements and actuator forces are thus confined to the axial elongation/compression motions and rates of the bipod legs.

The dynamic behaviour of the structure is described by the second order vector differential equation of the generalized coordinate vector \underline{q}

$$M\ddot{\underline{q}} + D\dot{\underline{q}} + S\underline{q} = \underline{B} \cdot \underline{f} \quad (2.1)$$

with $M = E^T M^* E = \text{diag} [1]$

$$D = E^T D^* E = \text{diag} [2 E_i \omega_i]$$

$$S = E^T S^* E = \text{diag} [\omega_i^2]$$

$$\underline{B} = E^T \underline{B}^* \quad (2.2)$$

and $\underline{q}^T = [q_1, \dots, q_{12}]$.

Thereby M^* , D^* and S^* are the mass, damping and stiffness matrices of the structure and B^* is the input force distribution matrix for the input force \underline{f} . These matrices are normalized with the eigenvector matrix E from the finite element analysis ¹⁾, which also yields the eigenfrequencies ω_i . The structural damping E_i is assumed to be 0.005 (0.5%) for $i = 1 \dots 12$.

Using only velocity measurements \underline{v} in the bipod legs, the measurement equation is determined as

$$\underline{v} = \underline{C} \cdot \dot{\underline{q}} \quad (2.3)$$

with

$$\underline{C} = \underline{C}^* \cdot \underline{E}.$$

Thereby C^* is the matrix containing the direction cosines of the bipod legs.

With the above nomenclature, the motion $\underline{x}(t)$ of the vertices 1-4 of the tetrahedral truss structure is given as

¹⁾ The data of the finite element analysis were delivered by K.R. Strunce from the Charles Stark Draper Laboratory, Cambridge, Mass., USA.

$$\underline{x}(t) = E \cdot \underline{q}(t) \quad (2.4)$$

with \underline{x} containing the ordered x, y and z components of the vertices.

2.2 Controller Design based on Dissipation Energy Maximization

Design Requirements

According to [3] the motion of the flexible truss structure specified by equations (2.1), (2.3) and (2.4) has to be controlled such that the following design requirements are met:

- 1) The physical model to be used in the controller design process consists of the first eight modes of the model. The modes 9 to 12 are to be treated as unmodelled modes, such that they cannot be used in any controller design process. However, they are to be used in the full-order evaluation model.
- 2) A minimum of 10 percent modal damping has to be assigned to the modes 1, 2, 4 and 5 and a minimum of 0,5 percent to the remaining modes of the 12 mode model. If that is achieved, the closed-loop response of the structure to certain initial conditions is such that the x and y components of the top vertex have magnitudes less than 0.0004 and 0.00025 units of length, respectively, after 20 seconds.

For fulfilling these design requirements the controller can employ all 6 actuators/sensors in the bipod legs or any reduced number of combinations of these actuators/sensors.

The primary objective of this model ¹⁾ is to provide a well-defined, non-trivial dynamical plant which serves as a common "test-bed" to illustrate, compare, and evaluate various controller design.

Controller Structure

A suitable low authority controller structure for vibration damping of the flexible truss structure is constant feedback of the velocities, measured in the bipod legs:

$$\underline{\dot{f}} = -K \cdot \underline{v} \quad (2.5)$$

Using these velocity measurements together with force actuators the closed loop dynamic equation of the structure is achieved by combining equation (2.1), (2.3) and (2.5):

$$M\ddot{\underline{q}} + [D+BKC] \dot{\underline{q}} + S\underline{q} = \underline{0} \quad (2.6)$$

For collocation (collocated actuators and sensors), the following equation is valid

$$B = C^T$$

If in addition K is symmetric and positive definite, closed loop stability can be guaranteed even for the unmodelled modes, as e.g. modes 9-12 of the truss structure. But fulfillment of the damping requirements is not included. Therefore a general feedback matrix K is chosen for the design process.

For the determination of this feedback matrix K fulfilling all design requirements, the following numerical optimization is performed:

Dissipation Energy Maximization

The sum of kinetic and potential energy in a mechanical system at an initial time t_0 is given as

$$W_0 = T + V = \frac{1}{2} \dot{\underline{q}}^T M \dot{\underline{q}} + \frac{1}{2} \underline{q}^T S \underline{q} \quad (2.7)$$

Differentiating with respect to time yields the energy rate

$$\dot{W}_0 = \frac{1}{2} \dot{\underline{q}}^T M \dot{\underline{q}} + \frac{1}{2} \dot{\underline{q}}^T M \dot{\underline{q}} + \frac{1}{2} \dot{\underline{q}}^T S \underline{q} + \frac{1}{2} \underline{q}^T S \dot{\underline{q}} = -\dot{\underline{q}}^T D \dot{\underline{q}} \quad (2.8)$$

extracted from the system by damping agencies.

¹⁾ In reference [3] in addition to a nominal model of the truss structure a modified model is specified. The controller has to fulfill the design requirements for both models.

Reintegration leads to

$$W_o = W_F = - \int_{t_o}^{\infty} \dot{\mathbf{q}}^T \mathbf{D} \dot{\mathbf{q}} dt \quad (2.9)$$

which is called dissipation energy of the system [5]. As \mathbf{D} is the matrix of modal damping, W_F represents the energy dissipated by internal velocity proportional friction.

Because of the absence of other dissipation factors, W_F is equal to the total initial energy W_o .

Now introducing output feedback according to equation (2.5) adds a new velocity proportional term \mathbf{BKC} to equation 2.6. This additional term splits into a symmetric dissipative part

$$\mathbf{D}_C = \frac{1}{2} [\mathbf{BKC} + (\mathbf{BKC})^T] \quad (2.10)$$

and a skew symmetric matrix of conservative forces

$$\mathbf{G}_C = \frac{1}{2} [\mathbf{BKC} - (\mathbf{BKC})^T] . \quad (2.11)$$

The matrix \mathbf{D}_C represents the damping induced by the velocity feedback and \mathbf{G}_C describes the gyroscopic effects of the feedback.

Analogue to equation (2.9), the integral

$$W_C = - \int_{t_o}^{\infty} \dot{\mathbf{q}}^T \mathbf{D}_C \dot{\mathbf{q}} dt \quad (2.12)$$

represents the energy which is dissipated by control action. The initial energy W_o of the system now partially is dissipated by friction and control

$$W_o = W_F + W_C . \quad (2.13)$$

The gyroscopic matrix \mathbf{G}_C thereby takes care of the "internal" energy transfer between different modes of the system, without changing the total energy of the system.

This dissipation energy W_C is used as the global optimization criterion, to determine the feedback matrix \mathbf{K} .

$$\max_{\mathbf{K}} W_C(\mathbf{K}) \rightarrow \mathbf{K}^* . \quad (2.14)$$

Additionally the damping requirements

$$\xi_1, \xi_2, \xi_4, \xi_5, \geq 10 \% \quad (2.15)$$

are formulated as constraints of the numerical optimization.

Thus the determination of the low authority controller is performed via numerical optimization, implementing the design requirements as constraints. This constraints may, besides damping, contain additional design requirements as shown in [12].

For this nonlinearly constrained optimization problem the recursive quadratic programming method of Han and Powell [6-9] explained in appendix A, is used efficiently. The numerical determination of the dissipation energy W_C is shown in appendix B.

2.3. Design Results

The maximization of the dissipation energy is performed for the system with 8 modes, as specified in the design requirements. For programming reasons additionally to the constraints of equation (2.15) $\xi_3 \geq 10\%$ was formulated. As initial state for optimization was chosen $\dot{q}_1(0), \dot{q}_2(0) = 2.0; \dot{q}_4(0), \dot{q}_5(0) = 0.2$ and zero elsewhere. Thus, large damping of modes 1, 2, 4 and 5 was expected.

The uncontrolled behaviour in the x, y plane of the line of sight (top vertex) after a typical disturbance is shown in Fig. 2.2.:

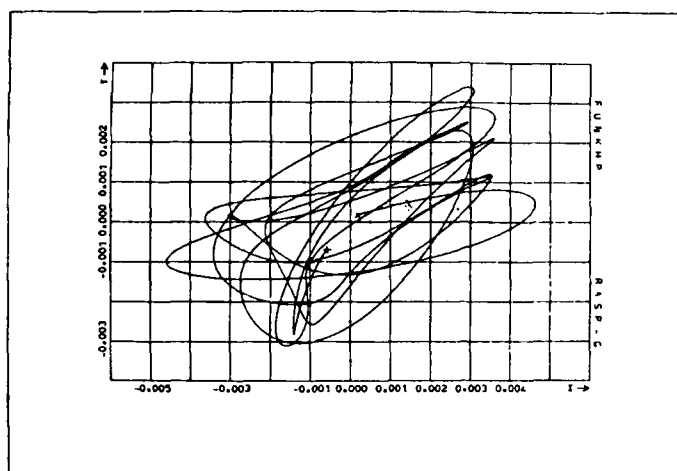


Figure 2.2 Motion of the line of sight without low authority control for 30s

Six actuators and sensors

The optimization using all six actuators and sensors yields the following feedback gain.

$$K = \begin{bmatrix} 14.68 & -1.83 & -13.58 & -11.71 & 8.70 & 4.44 \\ 5.29 & 7.86 & 12.07 & -13.67 & -0.83 & 10.71 \\ 10.53 & -7.81 & 17.00 & 2.91 & -13.82 & 0.43 \\ 10.48 & 13.43 & 15.98 & 16.11 & 11.28 & -15.99 \\ -8.61 & 10.60 & 14.53 & 14.59 & 16.49 & -13.98 \\ -15.10 & 5.37 & -15.99 & 9.55 & 4.76 & 16.14 \end{bmatrix} \quad (2.16)$$

This gain is the first during the path of optimization which fulfills the design requirements¹⁾. With continuing optimization, dissipation energy is increased using larger gains. As maximization of the dissipation energy is only a secondary means for changing the gains such that the design requirements are met, optimization is stopped, as soon as this is achieved.

Fig. 2.3 shows the line of sight motion using this feedback gain.

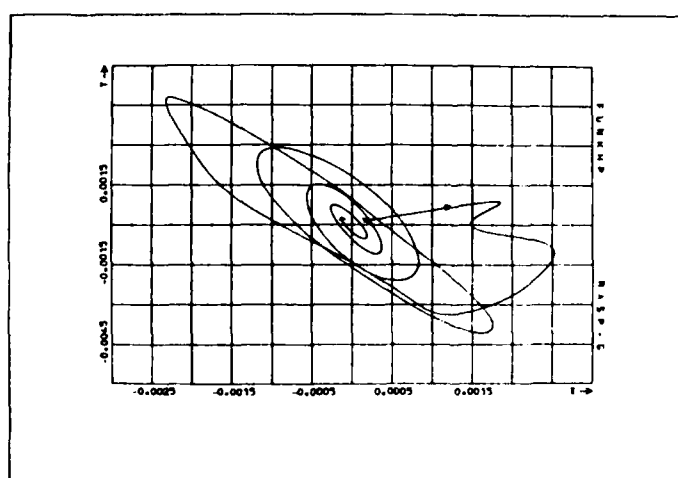


Figure 2.3 Motion of the line of sight using 6 actuators and sensor

¹⁾ This gain fulfills also the design requirements for a modified truss structure [12].

The next figures show the eigenvalue plots for the design model (Fig. 2.5) and the complete evaluation model (Fig. 2.4) with all 12 modes.

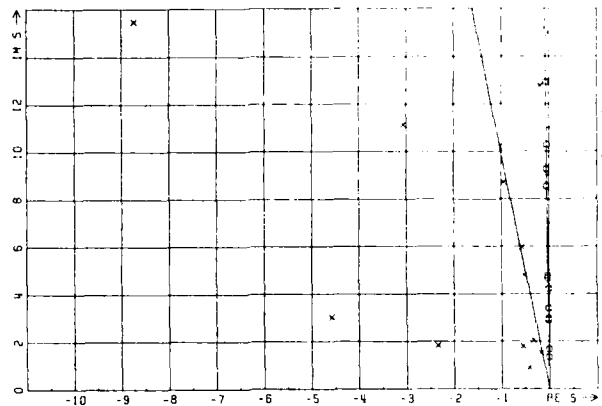


Figure 2.4 Eigenvalues of the evaluation model
(o open loop; x closed loop)

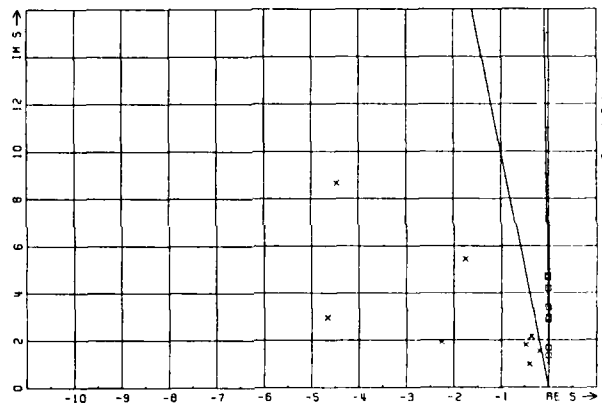


Figure 2.5 Eigenvalues of the design model
(o open loop; x closed loop)

Four actuators and sensors

Without using the sensors and actuators in the bipod legs of vertex 2 the design requirements could not be fulfilled. This is due to the fact, that modes 9 to 12 should not be considered in the design procedure. After 50 iterations the design conditions were almost met using very large feedback gains. However, this finally led to a decrease of the achieved damping for the complete evaluation model (spillover effect). Thus an intermediate solution was chosen which did not display large spillover effects.

Figures (2.6) to (2.8) present the achieved time behaviour and eigenvalue locations for the feedback gain

$$K = \begin{bmatrix} 43.64 & -16.12 & 26.68 & -0.63 \\ 30.83 & 31.40 & 31.12 & -4.22 \\ 8.03 & -29.98 & 8.01 & -37.23 \\ 16.00 & -36.47 & 6.36 & -3.91 \end{bmatrix} \quad (2.17)$$

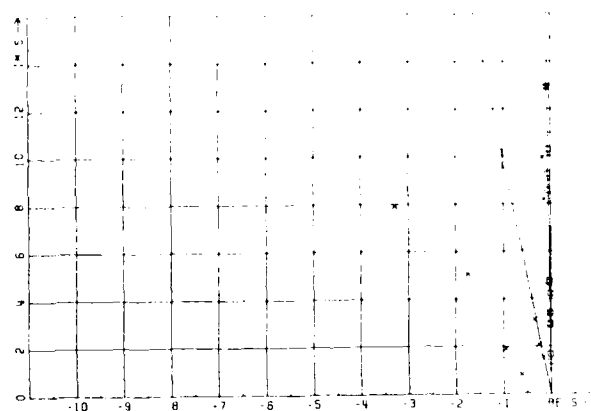


Figure 2.6 Eigenvalues of the evaluation model
(□ open loop; x closed loop)

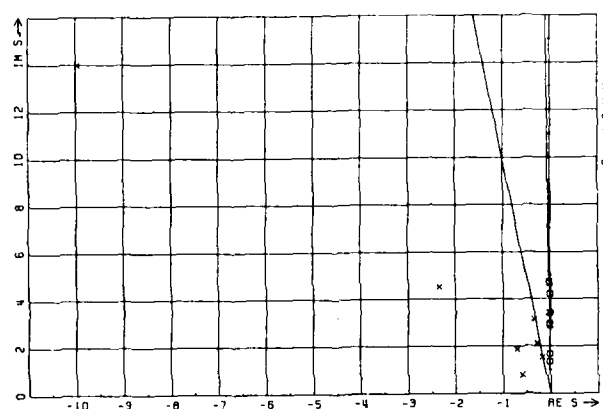


Figure 2.7 Eigenvalues of the design model
(□ open loop; closed loop)

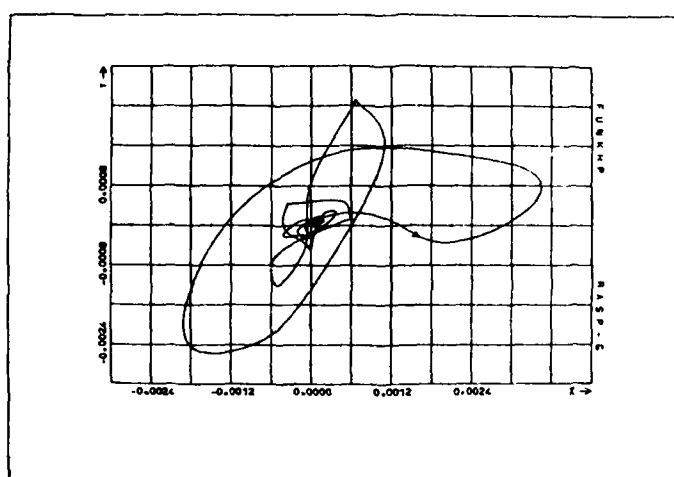


Figure 2.8 Motion of the line of sight using
4 actuators and sensors

3. Minimization of the Control Energy

Low authority controller design via minimization of the control energy is shown for the "Purdue Model", explained in the following:

3.1 "Purdue Model"

The model is a large two dimensional space vehicle which can be considered as a model of a solar power station. It is similar to a 5 km x 12.5 km plate with a central rigid body.

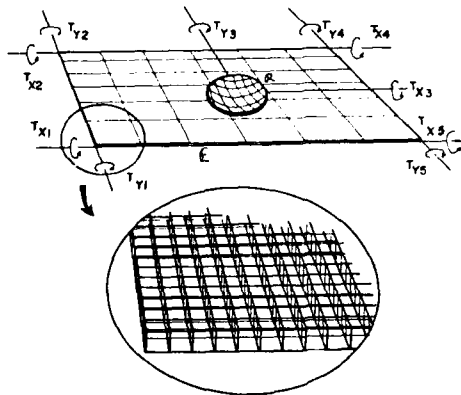


Figure 3.1 A large two-dimensional space vehicle

This plate was developed by Hablani [4] as a generic model of a large flexible space structure for control concept evaluation. At the four corners of the plate and in the center, x and y torquers are implemented. Collocated with these torquers are slope and rate sensors of the deflection. The structure of the model equations is identical to equations (2.1) and (2.3). The data of the relevant matrices are given in [4]. Contrary to the Draper Model this one uses torquers \underline{T} and angular velocity measurements $\underline{\dot{\theta}}$:

$$\begin{aligned} M\ddot{\underline{q}} + D\dot{\underline{q}} + S\underline{q} &= B \cdot \underline{T} \\ \underline{\dot{\theta}} &= C \cdot \dot{\underline{q}} \end{aligned} \quad (3.1)$$

The evaluation model contains eleven modes. The modal pattern of the extended model is displayed in appendix D. Structural damping of the open loop system is fixed to be 0.5% for all modes.

3.2 Controller Design based on Control Energy Minimization

Design Requirements

For the design of an attitude control system (high authority control) modal damping of the plate has to be augmented using low authority control.

Based on a low order design model the modal damping of the structure shall be increased from 0.5% to 5% using constant output feedback.

Model Reduction and Controller Structure

For development of the reduced order design model the model reduction technique of Litz [14] was applied. A reduced order model containing the flexible modes No. 5, 7, 8 and 11, according to appendix D, was considered adequate for controller design. The designed controllers shall be evaluated using the evaluation model with eleven modes.

As above, constant output feedback is chosen as controller structure

$$\underline{T} = -K \cdot \underline{\dot{\theta}}, \quad (3.2)$$

thus leading to the same closed loop equation as in section 2

$$M\ddot{\underline{q}} + [D + BKC] \dot{\underline{q}} + S\underline{q} = \underline{0}. \quad (3.3)$$

Control Energy Minimization

The experience gained with the dissipation energy maximization showed, that the large feedback gains may deteriorate the system behaviour because of spillover effects. Thus it is the objective to fulfil the design requirements with small feedback gains. As a means to achieve this objective, minimization of the control energy is chosen as performance criterion:

$$\min_K \int_{t_0}^{\infty} \underline{T}^T(t) \cdot \underline{T}(t) dt \rightarrow K^* \quad (3.4)$$

Additionally the damping requirements are formulated as constraints

$$\xi_i \geq 5\% \quad \text{for } i = 1, \dots, n \quad (3.5)$$

In the former case of dissipation energy maximization the optimization criterion was used as a means to change the gains such that the constraints are met. Augmenting the gains improved energy dissipation and required damping. The gradients of the criterion and of the constraints in general had the same sign. In the case of control energy minimization the gradients of the criterion and of the constraints have in general opposite signs. Thus augmenting the gains deteriorates the energy minimization but improves damping.

For performance of the optimization a modified version of the Han/Powell algorithm [6-9] using least squares approximation for the one-dimensional minimum search from Schittkowsky/Kraft [15] was used.

The control energy as function to be minimized is determined using the Liapunov equation in appendix B, but changing the matrix $D_c \rightarrow (KC)^T$. KC in the weighting matrix \tilde{Q} .

3.3. Design Results ¹⁾

For control energy minimization on initial state vector resulting from positive respectively negative torque impulses was chosen. Four different feedback strategies are investigated as shown in the following. Thereby only the torquers T_{x_1} , T_{y_1} , T_{x_3} , T_{y_3} , and T_{x_4} , T_{y_4} are used.

Local Feedback

The angular velocities measured in the lower left corner of the plate, i.e. $\dot{\theta}_{x_1}$ and $\dot{\theta}_{y_1}$, are fed back to the torquers T_{x_1} and T_{y_1} only. The same was done for the central angular velocities $\dot{\theta}_{x_3}$ and $\dot{\theta}_{y_3}$. They are fed back to the torquers T_{x_3} and T_{y_3} only. For the upper right corner it is vice versa. In this scheme the torquers use only information measured at the same location (local feedback!).

Optimization yield the following feedback gain:

$$K = \begin{bmatrix} 0.08 & -1.39 & & & & \\ 0.83 & -0.36 & & & & \\ & & 0.003 & -1.62 & & \\ & & 0.51 & 0.15 & & \\ & & & & -0.04 & -1.46 \\ 0 & & & & 0.79 & 0.71 \end{bmatrix} \quad (3.6)$$

Figures (3.2) and (3.3) show, that the design requirements are met for the design and also for the evaluation model.

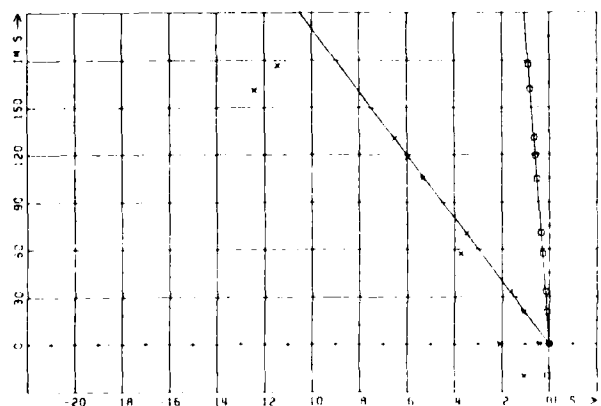


Figure 3.2 Eigenvalues of the evaluation model (\square open loop; \times closed loop)

¹⁾ These results were derived in the frame of the ESA-Study on "Investigation of the Attitude Control of Large Flexible Spacecraft" under ESTEC Contract No.5310/82/NL/BI, performed by DFVLR and Dornier-System.

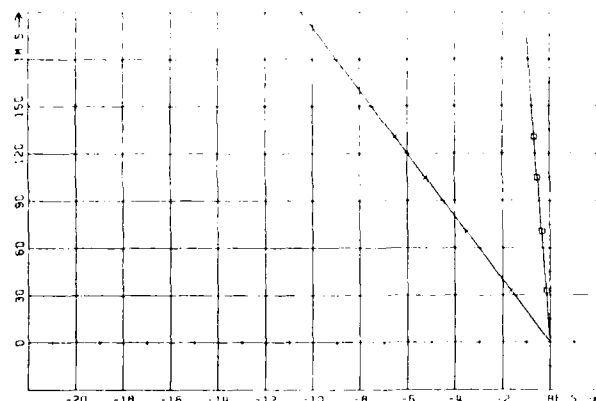


Figure 3.3 Eigenvalues of the design model
(□ open loop; x closed loop)

Axial Feedback

For this strategy the feedback gain matrix was restricted to be diagonal thus enforcing the following feedback scheme: $\dot{\theta}_{x_1} \rightarrow T_{x_1}$, $\dot{\theta}_{y_1} \rightarrow T_{y_1}$; $\dot{\theta}_{x_3} \rightarrow T_{x_3}$...

No stable solution was found for this feedback strategy. This is explained by the fact, that the dominating modes of the plates are the antimetric modes No. 5 and 7 (see appendix D).

Cross-Axial Feedback

For this strategy the following feedback scheme is used: $\dot{\theta}_{x_1} \rightarrow T_{y_1}$, $\dot{\theta}_{y_1} \rightarrow T_{x_1}$, $\dot{\theta}_{x_3} \rightarrow T_{y_3}$, Thus the measured angular velocities of the x-axes are locally fed back to the y-axes torquers and vice versa.

The following feedback matrix resulted:

$$K = \begin{bmatrix} 0 & -1.47 & & & & \\ 1.29 & 0 & & & & \\ \hline & & 0 & -1.09 & & \\ & & 0.185 & 0 & & \\ \hline & & & & 0 & -1.47 \\ 0 & & & & 1.29 & 0 \end{bmatrix} \quad (3.7)$$

The achieved eigenvalue distribution (see Fig. 3.4, 3.5) is excellent. This confirms that the dominant modes of the structure are the antimetric modes No. 5 and 7. A cross axial feedback scheme using only six feedback gains is sufficient to provide the required modal damping of 5%.

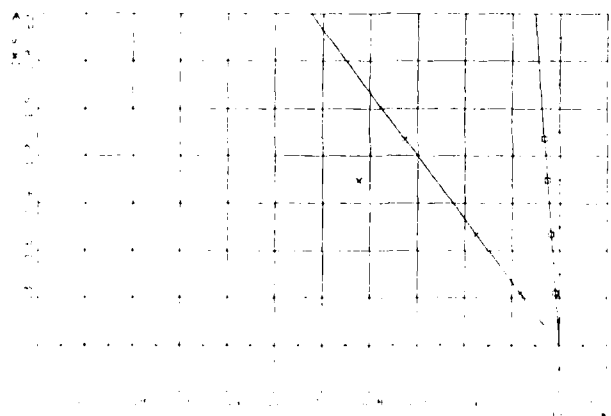


Figure 3.4 Eigenvalues of the design model
(□ open loop, x closed loop)

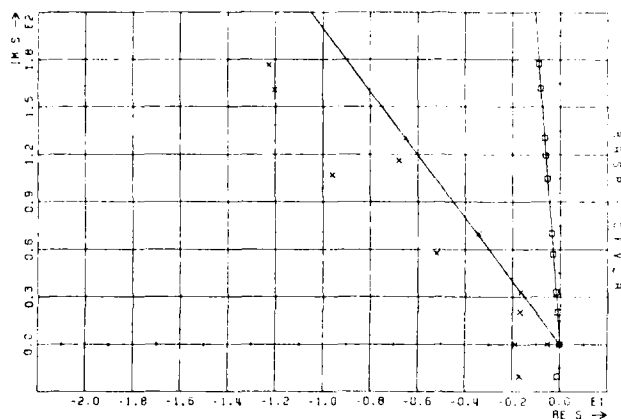


Figure 3.5 Eigenvalues of the evaluation model
(□ open loop, x closed loop)

Complete Feedback

Using the feedback gain matrix from local feedback as starting value the optimization routine yield the following feedback gain for the complete feedback case:

$$K = \begin{bmatrix} 0.016 & -1.60 & -0.004 & 0.22 & -0.07 & 0.017 \\ 0.70 & 0.13 & 0.17 & 0.12 & -0.07 & -0.039 \\ 0.014 & 0.07 & 0.015 & -1.65 & 0.01 & 0.045 \\ 0.085 & -0.012 & 0.51 & 0.01 & -0.003 & 0.06 \\ -0.048 & -0.031 & 0.0005 & 0.01 & -0.04 & -1.43 \\ -0.083 & -0.018 & 0.041 & 0.01 & 0.65 & 0.44 \end{bmatrix} \quad (3.8)$$

The eigenvalue locations Figs. (3.6) and (3.7) differ only slightly from those of local feedback, because the gains are not much different either.

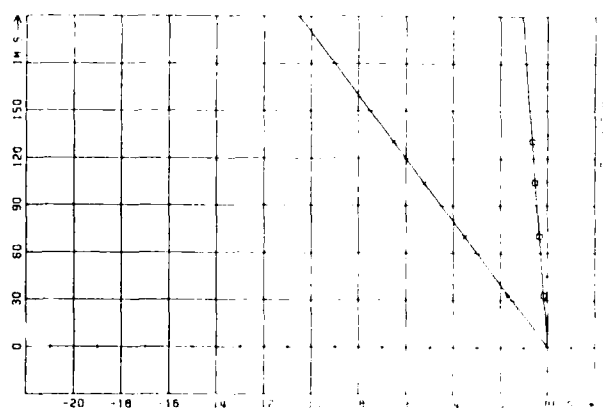


Figure 3.6 Eigenvalues of the design model
(□ open loop, x closed loop)

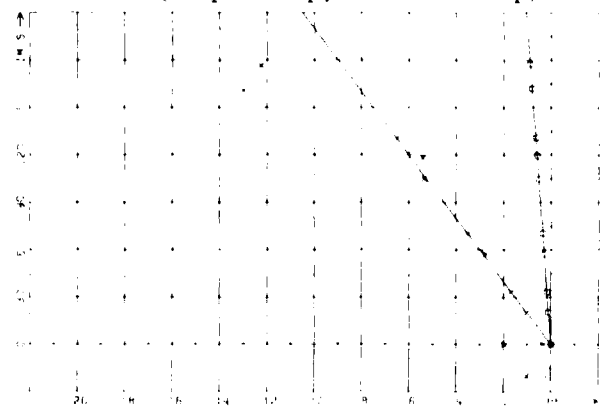


Figure 3.7 Eigenvalues of the evaluation model
(□ open loop, x closed loop)

4. Critical Assessment of the Design Methods

In the two proceeding chapters low authority controllers for flexible spacecraft were designed using numerical optimization methods. The given design requirements were fulfilled using different optimization criteria: first dissipation energy was maximized and second control energy was minimized. In both cases damping requirements were formulated as constraints.

As a result of dissipation energy maximization it could be shown, that this criterion finally led to large feedback gains. Thus the probability of unwanted spillover effects or even spillover instability increased.

However, minimizing the control energy, keeps feedback gains small. Thus the probability of spillover effects is reduced considerably, although its occurrence cannot be avoided completely.

Due to the highly efficient optimization routines used, the computation effort remains reasonable. Computation time for one optimization for a system with e.g. 4 modes, 4 inequality constraints and 16 gain elements to be optimized is about 15 sec CPU-time on an IBM main frame system. Thereby about 18 iteration steps with 300 function calls are performed. The computation time is used mainly for determination of the functional value. This is done for the above example by solving an 8-th order Liapunov equation (see appendix B).

Even for the Draper model with 8 modes, 10 inequality constraints and 36 gain elements to be optimized, a solution was found after about 120 sec CPU-time.

The advantage of low authority controller design by numerical optimization is its great variability for system design. Design requirements may not only be the damping requirements as above, but also robustness properties may be implemented [12] or actuators and sensors may be positioned on the flexible structure [2,5] or other design requirements may be postulated.

5. Conclusions

Due to the now existing high efficient software for nonlinear optimization with nonlinear constraints, optimization methods can be used advantageously in system design. This is shown for the design of low authority controllers for large flexible structures. Two different optimization criteria were implemented and the design requirements were formulated as constraints of the optimization. The main advantage of controller design using numerical optimization methods is the possibility of implementing arbitrary design requirements. Thus the use of optimization methods becomes a variable tool for system design even for high order systems.

6. References

- [1] Aubrun, J. N. Theory of the Control of Structures by Low Authority Controllers. J. Guidance and Control, vol 3, No. 5, Sept/Oct 1980, pp. 444-451.
- [2] Schulz, G. Dislocated Actuator/Sensor Positioning and Feedback Design for Flexible Structures. Appears in Journal of Guidance, Control and Dynamics, 1983.
- [3] Strunce, R.R. Final Report: Actively Controlled Structures Theory. Hegg, J.G. Vol. 1, Charles Stark Draper Laboratory Report, R 1538, Henderson, T.C. Dec 79. Lin, J.G.
- [4] Hablani, H.B. Generic Model of a Large Flexible Space Structure for Control Concept Evaluation. J. Guidance and Control, vol 4., No. 5, Sept/Oct. 1981, pp 558-561.
- [5] Schulz, G. Integrated Actuator/Sensor Positioning and Feedback Design for Large Flexible Structures. AIAA Guidance and Control Conf., Aug. 9-11, 1982, San Diego, Cal.
- [6] Powell, M.D.J. A Fast Algorithm for Nonlinearly Constrained Optimization Calculations. Numerical Analysis Conference, Dundee, Scotland, 1977; G.A. Watson (ed.) 1978, Springer Lecture Notes in Mathematics, No. 630.

- [7] Powell, M.D.J. The Convergence of Variable Metric Methods for Nonlinearly Constrained Optimization Calculations.
Nonlinear Programming Symposium 3, 1977, Madison, Wisconsin.
- [8] Han, S.P. A globally convergent method for nonlinear programming.
J. Optimization Theory and Application, 22, 1977, pp. 297-309.
- [9] Kraft, D., On the choice of Minimization Algorithms in Parametric Optimal Control Problems.
In "Optimization and Optimal Control", Lecture Notes in Control and Information Science, (A. Auslender, W. Oettli, J. Stoer, eds.), Springer, 1981, N.Y.
- [10] Brockett, R.W. Finite Dimensional Linear Systems.
John Wiley and Sons, Inc. N.Y. London, 1970.
- [11] Bartels, R.H. Solution of the Matrix Equation $AX + XB = C$.
Stewart, G.W. Communications of the ACM 1972, Vol. 15, No. 9, pp. 820-826.
- [12] Schulz, G. Robuste aktive Schwingungsdämpfung von Finite Elemente Strukturen in der Raumfahrt.
Heimbold, G. Zeitschrift für Flugwissenschaften und Weltraumforschung, Band 7, Heft 2, März/April 1983, Seite 91-99.
- [13] Hablani, H.B. Constrained and Unconstrained Modes: Some Modeling Aspects of Flexible Spacecraft.
Journal of Guidance and Control, vol. 5, No. 2, March-April 1982.
- [14] Litz, L. Practical Results by a New Modal Method of Order Reduction
Regelungstechnik 2.7, 1979.
- [15] Schittkowsky, K. Theorie und Anwendung der sequentiellen quadratischen Programmierung in Steuerungs- und Regelungsaufgaben.
Kraft, D. DFVLR-Forschungsbericht, 1983.

Appendix A

Numerical Optimization

The optimization problem of equation (2.14) and (2.15) represents a nonlinear optimization with constraints. The constraints are inequality constraints in E_1 . Han and Powell propose the following globally convergent extension of the damped Newton method for this problem [6-9]. Generate a sequence $\{y^k\}$ converging to a local solution y^* by means of recursively solving the following quadratic programming problem

$$\begin{aligned} \min_{\underline{d}} \quad & \{ \nabla F(y)^T \underline{d} + \frac{1}{2} \underline{d}^T H \underline{d} \} \\ \text{s.t.} \quad & \begin{cases} \nabla c_i(y)^T \underline{d} + c_i(y) = 0, \quad \forall i \in J_e, \\ \nabla c_i(y)^T \underline{d} + c_i(y) \geq 0, \quad \forall i \in J_i, \end{cases} \end{aligned}$$

which yields the search direction \underline{d} in which the next approximation to the solution is found:

$$y^{k+1} = y^k + \alpha_k \underline{d}^k.$$

The relaxation factor $\alpha_j \in (0,1]$ ensures global convergence; it results from a step-length algorithm by properly reducing the test function

$$R(y, \rho) = F(y) + \sum_{i \in J_e} \rho_i \cdot c_i(y) + \sum_{i \in J_i} \rho_i \cdot \min(c_i(y), 0),$$

with suitable chosen parameters ρ . H is an approximation to the Hessian of the Lagrangian

$$\mathcal{L}(\underline{y}, \underline{\lambda}) = F(\underline{y}) - \underline{\lambda}^T \underline{c}(\underline{y})$$

with F being the functional to be minimized. The algorithm belongs to the class of "variable metric methods for constrained optimization" and converges superlinearly [7].

Appendix B

Dissipation Energy Determination

Transforming equation (2.1) to state space form yields

$$\begin{bmatrix} \dot{\underline{q}} \\ \ddot{\underline{q}} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -M^{-1}S & -M^{-1}D \end{bmatrix} \begin{bmatrix} \underline{q} \\ \dot{\underline{q}} \end{bmatrix} + \begin{bmatrix} 0 \\ B \end{bmatrix} \cdot \underline{f}$$

$$\underline{v} = \begin{bmatrix} 0 & C \end{bmatrix} \begin{bmatrix} \underline{q} \\ \dot{\underline{q}} \end{bmatrix}$$

or in short form

$$\dot{\underline{z}} = A' \underline{z} + B' \underline{f}; \quad \underline{v} = C' \underline{z}$$

For the proposed output vector feedback, see equation (2.5), the closed loop stability matrix is given as

$$\tilde{A} = A' + B' K C' = \begin{bmatrix} 0 & I \\ -M^{-1}S & -M^{-1}(D + BKC) \end{bmatrix}$$

Using the definition

$$\tilde{Q} = \begin{bmatrix} 0 & 0 \\ 0 & D_c \end{bmatrix}$$

the controller induced dissipation energy (equation (2.12)) can be written as

$$W_c = - \int_{t_0}^{\infty} \dot{\underline{q}}^T D_c \dot{\underline{q}} dt = - \int_{t_0}^{\infty} [\underline{q}^T \dot{\underline{q}}^T] \tilde{Q} \begin{bmatrix} \underline{q} \\ \dot{\underline{q}} \end{bmatrix} dt \quad (*)$$

Applying standard state transformation techniques equation (*) yields

$$W_c = - \left[\underline{q}^T(0) \dot{\underline{q}}^T(0) \right] \int_{t_0}^{\infty} \tilde{e}^{A^T \cdot t} \tilde{Q} e^{\tilde{A} \cdot t} dt \begin{bmatrix} \underline{q}(0) \\ \dot{\underline{q}}(0) \end{bmatrix}$$

The solution of the integral expression defined as

$$P = - \int_{t_0}^{\infty} e^{\tilde{A}^T \cdot t} \tilde{Q} e^{\tilde{A} \cdot t} dt$$

is equivalent to the solution of the following Liapunov equation

$$\tilde{A}^T P + P \tilde{A} = \tilde{Q} \quad (**)$$

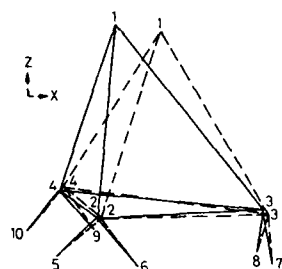
Thus the controller induced dissipation energy is determined as

$$W_c = - [q(0)^T \dot{q}(0)^T] P \begin{pmatrix} q(0) \\ \dot{q}(0) \end{pmatrix}$$

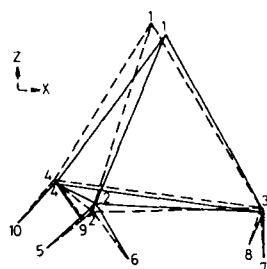
The determination of the dissipation energy W_c of a flexible structure with n modes requires the solution of Liapunov equation (**) for $2n \times 2n$ matrices. The solution of equation (**) exists and is unique as long as the eigenvalues of \tilde{A} have negative real parts [10]. The numerical solution of the Liapunov equation is done efficiently using the algorithm of Bartels and Stewart, as symmetry of \tilde{Q} is guaranteed [11]. During the optimization, its path is restricted to matrices \tilde{A} having eigenvalues with negative real parts.

Appendix C

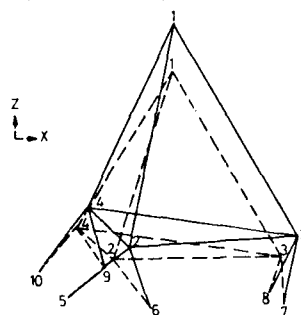
Mode Shapes of the Draper Model (taken from the original report [3])



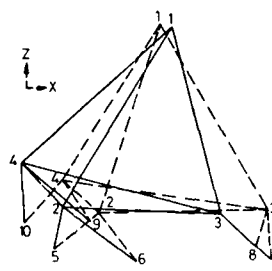
MODE 1
FREQUENCY = 0,214 Hz



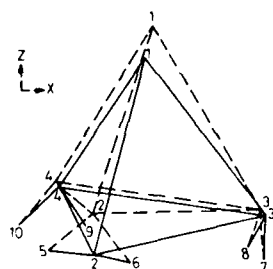
MODE 2
FREQUENCY = 0,265 Hz



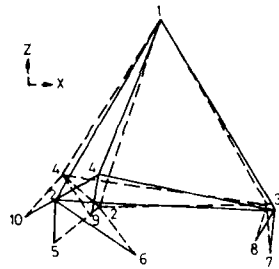
MODE 3
FREQUENCY = 0,460 Hz



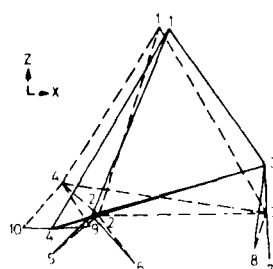
MODE 4
FREQUENCY = 0,471 Hz



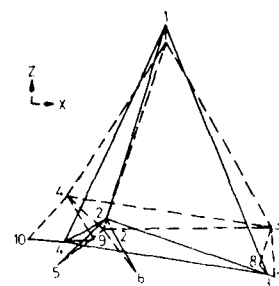
MODE 5
FREQUENCY = 0,541 Hz



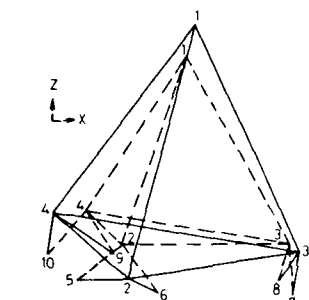
MODE 6
FREQUENCY = 0,669 Hz



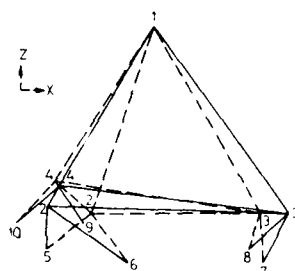
MODE 7
FREQUENCY = 0,742 Hz



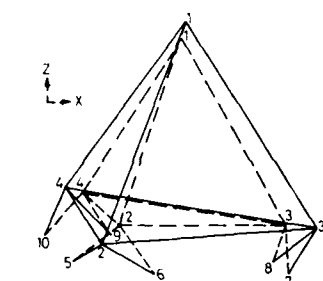
MODE 8
FREQUENCY = 0,777 Hz



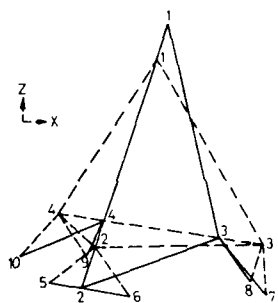
MODE 9
FREQUENCY = 1,357 Hz



MODE 10
FREQUENCY = 1,471 Hz



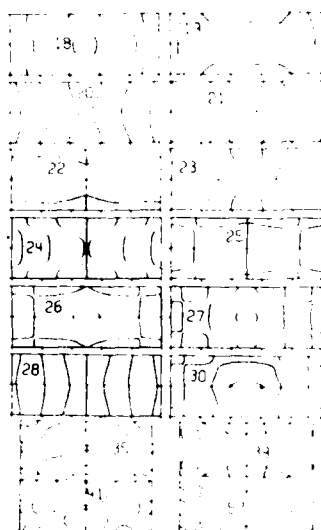
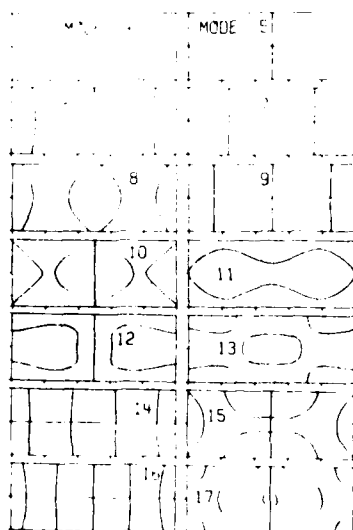
MODE 12
FREQUENCY = 2,054 Hz



MODE 11
FREQUENCY = 1,637 Hz

Appendix D

Nodal Pattern of the Purdue Model (taken from [13])



Achille Danesi

$$FF = (f_1 + f_2 + \dots + f_n)$$

A new strategy in controlling the modal shapes of large structure in space is presented in this study. An active low authority modal control system, consisting of a discrete number of servo system units (M.C.U.) distributed along flexible masts supporting RF radiators, is provided to measure and control the local structural deformations in order to obtain a modal shapes resulting in acceptable pointing error for the RF radiators. Each servo unit is conceived as a model following control system implementing a strategy based on F.F.T. pairs computations. The control efforts applied by the M.C.U.s to the mast control points are made proportional to the spectral error function generated by a real time F.F.T. dedicated microprocessor as the difference between the measured and desired modal spectra relative to the antenna angular deflections in respect to a fixed set point. The decentralized control system organized by a central controller residing in the master M.C.U., will force the antenna to match the model dynamical behaviour which is expected to improve the antenna pointing accuracy within the L.S.S. design requirements.

A control process correlated with the spectral requirements becomes a stringent necessity particularly in aerospace applications where the system dynamic behaviour must be essentially regulated at a particular frequency range established by the designer. The need of a control strategy based on the frequency domain data information is felt particularly in the research field treating the structural modal control based on the natural frequencies identification processes. A control strategy based on spectral data computed by a real time high speed F.F.T. dedicated microprocessor has been presented in Ref.1, treating the analysis and synthesis of a digital feedback control system applying the new concepts of the frequency domain model following strategy; in the proposed model following configuration the feedback error in the correspondent analog counterpart is replaced by the error function obtained as the inverse F.F.T. transform of the spectral difference between the system and model response, the last one stored in a programmable microprocessor memory area. The spectra comparison is performed in the same time slot in which the system response spectral lines are computed from the same number of time observations. The total time required to carry out the F.F.T. and successive I.F.F.T. computations, the last one employing, with only minor changes, the same algorithm used for the F.F.T., has to be less then the controlled system settling time in order to obtain a regulation process compatible with the time domain system performances; such a restriction imposes a limit on the processor computing speed. The feasibility of such control system comes out from the enormous progress in microprocessor technology allowing a drastic reduction in the time required to perform I/O and basic arithmetic operations involved in the F.F.T. algorithm computations. In consideration of the processor technology state of art, the processor computing speed doesn't constitute a problem for the class of space application treated.

The model, defined in the frequency domain, is programmed to comply with the frequency constraints predicted for the system. The feedback process developed in the servo loop will force the system to respond as the model, "reflecting", to an approximation depending on the designer ingenuity, the frequency characteristics imposed on the model. The main factors affecting the degree of system convergence toward the model are the spectral line selectivity, which depends on the process sampling rate, the F.F.T. algorithm numerical properties and associated software employed to implement it. In order to improve the convergence, the error spectral function can be, before going into the I.F.F.T. process, modified by means of digital filtering; in principle using appropriate F.I.R. or I.K.R. algorithms, each spectrum line may be corrected in the range of the frequencies considered in the F.F.T. computation.

The system time response is generally expected to differ somehow from the time response correspondent to the model spectrum; this effect, due to some degree of incompatibility of system and model spectral contents, may be adapted to the design requirements reaching a reasonable compromise between the time and frequency domain specifications by

choosing a proper digital filtering algorithm.

The main advantages arising from the control strategy based on the F.F.T. data informations can be summarized as follows:

- the system frequency spectrum can be observed in a single or repetitive time intervals, the length of which can be chosen on the basis of the system bandwidth and F.F.T. frequency selectivity
- a number of model spectra can be programmed and stored in the microprocessor memories to obtain a set of frequency constraints to be reflected on system time response
- the control system is structurally suited to implement a digital filtering process, as may be requested to solve advanced control algorithms
- a new potential authority in noise rejection can be achieved as result of an high resolution obtainable in single spectral line measurement
- the spectral power density informations, computed on the available F.F.T. data, can be directly used to implement a control strategy tending to regulate the system inherent spectral energy.

A servo unit, as it has been considered in Ref.1, has been employed as a Modal Control Unit (M.C.U.) in a decentralized active modal control of a flexible masts supporting large communication antennas in a L.S.S. platform sketched in Fig. 1.

The main objective proposed in the present work is to control the antenna angular displacements in order to improve its L.O.S. pointing accuracy.

The description of the structural model assumed for the L.S.S. platform and its modal characteristics are given respectively in Sect. 2 and Sect.3. The concepts of the proposed modal control strategy with its physical and analytical aspects are discussed in Sect.4. In Section 5 some details of system design and implementation are treated. Simulation results are presented in the last section.

2 - Platform structural model

The structural configuration assumed for the L.S.S. platform is depicted in Fig.2. It consists of a symmetrical structure with a large hub and platform arms supporting the solar panels and payloads. A 45 meters diameter communication antennas are attached to the mast by rigid mounting pads. The tip to tip dimension of the supporting masts is fixed at 33.75 m.; the antenna mass, which includes the mounting pad, feed unit and electrical connections, amounts to 434.209 Kg. The mast structure consists of beam like lattices having triangular cross sections providing a minimum number of longitudinal members; five identical bay elements with 7.5 m. length longitudinal bars, 5 mt. long battens and reinforcing diagonal stringers, are the constituent parts of the platform arms supporting the antennas. Considering a mass density of 2768 Kg/m³, the mass of the single bay is estimated at the value of 13.469 Kg. The stiffness product (E.I) taken into consideration is 1.57 10⁶ N.m².

In the present study only the platform arms are considered flexible while all the other platform structural components, including the antenna dish and feed, are assumed rigid. Since the investigation afforded is essentially directed to solve a dynamical control problem, only a relatively small number of degrees of freedom, correspondent to the lowest natural frequencies, will be sufficient, with good engineering accuracy, to analyze the dynamical behaviour of the antennas involved in the flexibility effects of the supporting masts. For that reason a lumped mass structural model, where the mass of each platform arm is reduced to a discrete number of rigid masses connected by weightless connectors having the same elastic properties as the physical structure, has been considered for the modal analysis. In conformity with the assumed structural configuration, each platform arms is reduced to five discrete masses, each concentrated at the center of mass of the bay elements. Since the system is symmetrical in respect to the central hub, only half of it needs to be considered for the analysis.

To solve the structural dynamic problem, the common modal analysis approach has been employed. The natural frequencies for the 5 D.O.F. undamped structural model subjected only to the bending elastic deformation, spans from the value of 0.18945 Hz. (first bending mode natural frequency) to 23.293 Hz. (5-th bending mode natural frequency); The first lowest frequencies are dominant in beam elastic behaviour.

3 - The frequency domain model following concepts

In order to illustrate the strategy involved in the modal control process investigated in this study, consider the second order differential matrix equation describing the coupled forced dynamics of the assumed lumped masses model:

$$[M] \ddot{q}(t) + [K] q(t) = F(t) \quad (1)$$

where $[M]$ and $[K]$ are respectively the symmetrical mass and stiffness matrices, q is the vector of the generalized coordinates and F is a vector of the external forces applied to the grid points. For the sake of simplicity, the generalized coordinates are referred to the angular deflection of the idealized beam sections by defining the state vector $x(t)$ the element of which are, for each degree of freedom, the angular beam deflections and their first derivatives. The equation (1) is then transformed in a set of first order differential equations:

$$\dot{x}(t) = A x(t) + B r(t) \quad (2)$$

where A and B are respectively the state and control matrices with elements expressed in terms of masses and stiffness coefficients. The dynamical behaviour of an established beam reference section is obtained integrating, with congruent initial conditions, the state equation (2) and selecting, as output variable, the angular displacement observed at the reference section by means of the scalar output equation:

$$a_r(t) = C x(t)$$

This time function is measured, at discrete time intervals, by an angular displacement sensor installed at the reference section; the output of this sensor is Fourier transformed and the resulting normalized discrete frequency spectrum (D.F.T.) is expressed by:

$$A_r(k) = \sum_{n=0}^{N-1} a_r(n) e^{-j(\frac{2\pi}{N})k} \quad k = 0, 1, 2, \dots, N-1 \quad (3)$$

where N is the number of the angular displacement (a_r) samples, assumed equal to the correspondent number of the computed frequency samples. This spectrum is observed in a specified time slot, established in function of the system bandwidth.

Consider a time response model $a_m(t)$ for the reference section angular displacements which can be generated by solving, in a digital filter, the following N -order difference equation:

$$a_m(t) = \sum_{n=0}^N b_i r(n-i) - \sum_{i=1}^N a_i a_m(n-i) \quad (4)$$

relating the current output samples to the past output samples, stored in computer memories, and the present and past input samples.

The discretized input forcing function $r(n)$ is supposed to be applied simultaneously to the system (1) and model generator (4) inputs.

The D.F.T. spectrum for the response model (4) is obtained as:

$$A_m(k) = \sum_{n=0}^{N-1} a_m(n) e^{-j(\frac{2\pi}{N})k} \quad k = 0, 1, 2, \dots, N-1 \quad (5)$$

The spectral error function $E(k)$ correspondent to the time error function:

$$e(t) = a_m(t) - a_r(t) \quad (6)$$

will be expressed by:

$$E(k) = A_m(k) - A_r(k) = \sum_{n=0}^{N-1} [a_m(n) - a_r(n)] e^{-j(\frac{2\pi}{N})k} \quad k = 0, 1, 2, \dots, N-1 \quad (7)$$

Forcing to a null value the spectral error (7) in a frequency band of interest means, in principle, to impose to the controlled system (1) the frequency constraints established for the model spectrum (5). This theoretical result can be approximated implementing a discrete feedback process driving a servosystem in its effort to match the system and model spectra. Since it is desirable to maintain a conventional analog configuration for such feedback model following structure, the spectral error function (7), must be transformed back in the time domain applying the inverse discrete Fourier transform (I.D.F.T.) process to the function $E(k)$:

$$e(n) = \frac{1}{N} \sum_{k=0}^{N-1} E(k) e^{j(\frac{2\pi}{N})k} \quad n = 0, 1, \dots, N-1 \quad (8)$$

The correspondent continuous error function (6) is obtained smoothing the discrete function (8).

Since it is advisable to perform the spectral computational processes in an on-line computer located in the feedback branch of a conventional servo loop, the model-following configuration shown in Fig.3 has been adopted. The model error defined:

$$e_m(t) = r(t) - a_m(t) \quad (9)$$

is generated solving the first order state equation:

$$\dot{e}_m(t) = A_e e_m(t) + B_e r(t) \quad (10)$$

$$e_m(t) = C e_m(t)$$

The feedback branch is carrying the following information:

$$f(t) = e_m(t) - a_r(t) \quad (11)$$

yielding, at the servo unit input, the error function:

$$e(t) = r(t) - f(t) = a_m(t) - a_r(t) \quad (12)$$

ready defined in (6). Since all the data flowing on the feedback branch are the results of the digital processing carried out by a digital computer, the function $f(t)$ as a matter of fact is discrete in nature and comes out from a D.F.T. process:

$$F(k) = E_m(k) + A_r(k) \quad (13)$$

where $A_r(k)$ was defined in (3) and $E_m(k)$ is the D.F.T. of the discrete error function $e_m(n)$ given in (9). The continuous function $f(t)$ is obtained by a smoothing process applied to the result of the discrete inverse transformation (I.D.F.T.) of the computed spectrum $F(k)$:

$$f(n) = \frac{1}{N} \sum_{k=0}^{N-1} F(k) e^{j\left(\frac{2\pi}{N}\right)kn} \quad (14)$$

$$n = 0, 1, 2, \dots, N-1$$

Generally before going into the I.D.F.T. process, the feedback spectrum $F(k)$ is digitally processed in order to impose on it some frequency window functions or to include corrections in the attempt to improve the system and model responses matching. The modified feedback spectrum will be obtained as a result of the digital process expressed in z-domain:

$$F_c(z) = G(z) F(z) \quad (15)$$

where the transfer function $G(z)$ can be generally expressed in M-polynomial form:

$$G(z) = \frac{\sum_{i=1}^M a_i z^{-i}}{\sum_{i=0}^M b_i z^{-i}} \quad (16)$$

4 - The spectral modal control

The frequency domain model following control principles sketched in Fig.4 are applied to implement a servo-unit which is a part of a organized active modal control system; it has been conceived as a compact device the constituent parts of which are an optical angular sensor, a force actuator and, for the master unit, a microprocessor. As indicated in Fig.5, the servo unit is configured as a feedback structure where the computing unit with its interfaces and the servoactuator with its power amplifier are respectively located in the feedback and forward path. A discrete number of such servo unit are distributed along the mast length constituting a decentralized active control system generating forces between parts of the mast structure. The servo control efforts are regulated in order to force the mast reference section, where the antenna is attached, to reproduce the model dynamical behaviour as predicted by the designer. To make clear this point suppose to choose a model with a stiffness product (EI) greater than the value relative to the actual mast and to apply to the last a force distribution such that the dynamical response of the actual mast becomes coincident with that predicted for the model; since the model has a greater stiffness it is reasonable to expect a lower angular deformation of the mast section assumed as reference for the active mode control. Taking into consideration several practical limiting factors such the maximum efforts level available for the control purposes, the size and the weight of the actuators and the interacting struc-

tural coupling effects between the mast sections supporting the servo units, only a partial matching must be expected. The present work is directed to investigate at this regard the effectiveness of the proposed modal control system in terms of reduction of the reference mast section angular displacement and to prove that it is large enough to improve significantly the antenna pointing accuracy. The design objective afforded in the following is to provide a control strategy regulating, in resemblance to a specified dynamical model, the angular deflections of the reference section. All the servo units which are part of the active control structure, and for that reason referred as Modal Control Units (M.C.U.s) concur in the generation of the force and moment distribution along the mast, as requested to provide the best obtainable system-model matching.

Since it is advisable to perform the required spectral computations involved in the modal control in proximity of the optical angular displacement sensor, which is expected to be integrated with the mast structure in the reference section location, the computing unit will reside in the M.C.U. linked to the reference section.

The control efforts generated by the other M.C.U.s are developed in coordination to the data flow originated in the computing unit; furthermore, in order to obtain the desired force distribution along the mast, different weighting gains must be programmed and assigned to the various M.C.U.s. For these reasons the M.C.U. linked to the mast reference section, on the ground of its memory and computing capability, will operate as centralized controller making the proposed active modal control (A.M.C.) system, as depicted in Fig.5 a decentralized control organized by a central controller residing in a "Master" M.C.U.

Continuing the analitic approach given in the preceding section, the force applied by each M.C.U. actuator to the correspondent grid point is made proportional to the time error function $e(t)$ given in (12) and its first derivative. Considering zero the external forcing function $r(t)$ applied to the model-following control structure and referring to (11) and (12), it can be proved that:

$$e(t) = a_m(t) - a_r(t) = -f(t) \quad (17)$$

The force distribution along the mast can be expressed as a vector:

$$P_i(t) = -K \cdot y(t) \quad (18)$$

$$i = 1, 2, \dots, N$$

where $y(t)$ is the excitation vector:

$$y(t) = \begin{bmatrix} f(t) \\ \dot{f}(t) \end{bmatrix} \quad (19)$$

and K is a gain matrix the element of which are the proportional (k_p) and derivative (k_d) gains implementing, for each M.C.U., a P.D. controller.

To save computational time consuming, a Fast Fourier Transform (F.F.T.) algorithm has been used in the D.F.T. computation; since the discrete Fourier transform is an exactly reversible process, the same F.F.T. algorithm has been used for the I.D.F.T. computation.

The F.F.T. algorithm has been widely treated in the literature; here few notes are given to introduce the F.F.T. implementation in a real time dedicated microprocessor. To provide a good trade off in computing speed and hardware complexity, the 2-radix decimation in time algorithm is taken into consideration. This algorithm is defined for N time and N frequency points, where N is a power of two. Breaking systematically the N points D.F.T. into two $N/2$ points D.F.T., the original D.F.T. results in a collection of two points D.F.T. allowing the maximum reduction in the total number of elementary operations involved in the algorithm. The frequency spectrum of a time function is computed by a repetitive use of a computational block (butterfly) the input of which is arranged as a sequence of time samples referred as even and odd data points. Each butterfly solves an algorithm, given for the k -th spectrum line, by the expression:

$$F(k) = F_1(k) + W^k F_2(k) \quad (20)$$

where the constituent spectra are defined:

$$F_1(k) = \sum_{n=0}^{N/2-1} f(2nT) \bar{W}^{nk}$$

$$F_2(k) = \sum_{n=0}^{N/2-1} f(2nT+T) \bar{W}^{nk}$$

$$n=k = 0, 1, \dots, N/2-1$$

where $f(n)$ is the observed time function and T is the sampling time. Since the phase factor W and \bar{W} are complex quantities, the frequency functions $F_1(k)$ and $F_2(k)$ are complex functions. Referring specifically to the feedback spectrum $F(K)$, it can be put in

the form:

$$F(k) = F.F.T. \left[f(n) \right] = M_F(k) e^{j\varphi_F(k)} \quad (21)$$

where the feedback spectrum is expressed in terms of the module $M_F(k)$ and phase $\varphi_F(k)$ resulting from the F.F.T. computations. A digital filtering process are generally applied to the module spectrum $M_F(k)$ to correct it by the phase effects in the range of frequencies which are judged to be dominant in the particular application at hand. Furthermore a digital process may be required to apply particular frequency windows to the computed spectrum to avoid spectral errors across the band and smearing at the band edges which can result in poor frequency selectivity; F.I.R. filtering process may be also required in order to improve, if required, the system - model matching.

Translating the corrected spectrum module in the time domain by means of the I.F.F.T. process, the discrete feedback function $f(n)$ is obtained. Employing the same digital filter used for spectral corrections, the numerical derivative of the discrete function $f(n)$ is computed solving the first order difference equation involving the present and past values of the computed $f(n)$. Smoothing out these numerical results, the excitation vector (19) is obtained. The actuating signals for the M.C.U.s actuators are generated applying to the computed numerical data the appropriate weighting coefficients which are established taking into account the different M.C.U. authority in generating the control moments. This regulation process is implemented multiplying the computed discrete function $f(n)$ by a different set of numerical values established for the proportional and derivative gains, collected in data vectors and stored in the computer memories; the resulting actuating data are addressed, through a multiplexer and smoothing units, to the respective M.C.U. input.

5 - Master M.C.U. implementation

The master M.C.U. is implemented as an electro-optical system integrated with the supporting structure on which the active control efforts are applied; it consists, of an optical sensor with its electronic back-up, a microprocessor which has a function of central processor for the decentralized modal control units distributed along the mast and a force actuator developing the active force excitation. In Fig.6, the systemistic scheme of the master modal unit, as it has been in principle conceived, is given. On the basis of knowledge of the general characteristics of the existing engineering products offered by the present aerospace technology, this modal control package is expected to have a mass which is a very few percent of the supporting mast bay.

In the following the general informations on the master M.C.U.'s constituent components are given.

- The electro-optical sensor

To measure the angular rotation of the assumed mast reference section due to bending deformations, an optical sensor is, in terms of resolution, reliability and installation flexibility, the most convenient for L.S.S. applications. The sensor taken into consideration in this study, tending mainly to investigate the system feasibility, is a particular application of the device presented in Ref. 2 as a position encoding sensor for static shapes measurements in L.S.S. structures. Here this kind of sensor is adapted to measure the angular displacement of the mast reference section in respect to a specified set point. This device uses a coaxial transmitter-receiver pulsed laser diode sensor with a sweeping capability in a specified angular range. By means of a photosensitive detector the angle (β_m) of the laser beam streaking a reflector point in the mast section identifying the observed reference section in respect to the laser optical axis and the distance (d_m) of the reflector from the collimator laser lens, are measured. The basic principle in the measurement process is that the emitted and returned radiation from the reflector detected by the laser head are, with a time delay introduced by a fiber optic unit, imaged on a photosensitive array where the relative coordinates are observed and measured. The signal at the photosensitive device output is processed by a microprocessor where the geometrical relation the measured quantities β_m and d_m and the angle (φ) through which the supposed untwisted reference section has rotated in respect to a nominal reference plane, is solved for the last unknown quantity. If this arrangement includes a discrete number of reflectors on the reference section the angular measurements can be corrected by the twisting effects on the observed section. This electro-optical device, which is expected to have, in the structural shape measurements, an accuracy in the order of 0,2±0,003 mm., when used in the angular measurements may have a resolution in the order of 0,1±2 arc-second which is considered satisfactory for the L.S.S. applications.

- The central processor unit

The central processor unit in the modal control system is an high speed-real time microprocessor implemented with multiplier-accumulator units offered, Ref. 3, by the present advanced technology which allow to build an extremely efficient data arithmetic

unit; this, when supported an appropriate control unit, can solve time and frequency domain algorithms like those involved in digital filtering and spectral analysis.

The modal control process developed by the central processor is based on the following computing stages:

- temporary storage in the computer dynamic memory area of the angular displacement data delivered, through an A/D converter, by the electro-optical sensor
- Computation of the angular displacement spectrum $A_r(k)$
- Extraction from the ROM memory area of the stored model error spectrum $E_m(k)$
- Computation of the feedback spectrum $F(k)$ as a sum of the two spectra $A_r(k)$ and $E_m(k)$
- Corrections applied to the feedback spectrum $F(k)$ by digital filtering
- Computation of the discrete feedback function $f(n)$ by means of the I.F.F.T. algorithm application
- Computation of numerical derivative of the discrete feedback function $f(n)$
- Multiplication of $f(n)$ by the gain matrix elements stored in the ROM area
- Selection of the output control vectors to be routed, via external multiplexing, to the M.C.U.s inputs.

- Force Servo actuators

To provide the distributed active actuations in the modal control process, D.C. force actuators, each capable of exerting a force up to 2.24 N. with a resolution of 0.01 N. has been considered. The intensity of forces, applied to the center of mass of the mast section by the servo actuator supported by the structural elements of the preceding bay, depends on the excitation level established by the central processor.

6 - Slaved M.C.U. implementation

The slaved M.C.U. is simply implemented with an electrical linear displacement pick-up connected to the bay's longitudinal bar, a force actuator and the common devices included in a feedback servosystem.

7 - Design parameters specification

To plan the active control process based on 2-radix decimation in time F.F.T. algorithm, specifications on the following design parameters are required:

- Spectrum recurrence frequency (f_s) - For a general band limited complex function, the correspondent discrete Fourier transform yields a sequence of periodic spectra at frequency f_s
- Spectrum bandwidth BW - Each spectrum will span a frequency range the length of which is the frequency bandwidth relative to the frequency contents of the transformed time function. To cover the full signal bandwidth $(BW)_s$ must be:

$$BW = (BW)_s$$

- Number of points in the F.F.T. spectrum (NF) - The spectrum at the processor output will appear as an NF-points discrete function. The 2-radix F.F.T. algorithm is based on a number of spectral points which is a power of two:

$$NF = 2^P \quad (P = 1, 2, \dots) \quad (22)$$

- Frequency resolution (R_F) - It is defined as the minimum frequency separation between two consecutive spectrum lines:

$$R_F = \frac{BW}{NF} \quad (23)$$

- Sampling time (T_s) - The time interval corresponding to sampling process on the observed time function is related to the spectrum bandwidth by:

$$T_s = \frac{2\pi}{BW} \quad (24)$$

The sampling frequency:

$$f_s = \frac{1}{T_s}$$

must be verified for the Nyquist criteria to avoid aliasing and ripple effects on the computed F.F.T. spectrum

- Number of time function samples (NT) - In order to accommodate the discrete nature of the Fourier transform and to obtain significant advantages in the processor hardware

and software, the number of samples for the the time function to be transformed is set equal to the number of points in the correspondent F.F.T. spectrum:

$$NT = NF = M \quad (25)$$

- Time function observation time (S_t) - The time function to be F.F.T. transformed is observed in a single o repetitive time slots lasting:

$$S_t = M \cdot T_s \quad (26)$$

For the model following active control design, the F.F.T. spectrum bandwidth and resolution for a N-degree of freedom structure are respectively referred to the highest and lowest natural frequencies in the range including the actual (A) and model (M) model frequencies:

$$BW \geq \max_{(A,M)} (f_n) \quad N=N_{\max} \quad (27)$$

$$R_F \leq \min_{(S,M)} (f_n) \quad N=N_{\min} \quad (28)$$

In the present study the model spectrum has been chosen referring to an ideal structure having the same geometric and mass distribution of the actual mast structure but with greater stiffness product (EI). The natural frequencies (f_n) for the proposed model structure are shown in Table I where, for comparison purpose, the correspondent values for the actual structure are also indicated. In the same table the values of the linear (δ_s^R) and angular ($\alpha_{f_{n1}}^R$) static displacements relative to the mast reference section due^s to the unit force ($\frac{1}{2}$ N.) applied to the mast tip, are given. In Table 2 the D.F.T. data adopted in the application treated in this study are specified

TABLE I - Model elastic characteristics

Datum	Dim.	Model Struct. EI = $3.57 \cdot 10^6$ (Nm ²)	Actual Struct. EI = $1.57 \cdot 10^6$ (N.m ²)
δ_s^R	m.	$1.29108 \cdot 10^{-3}$	$2.9377 \cdot 10^{-3}$
$\alpha_{f_{n1}}^R$	deg.	$8.5122 \cdot 10^{-3}$	$19.4851 \cdot 10^{-3}$
f_{n1}	Hertz	0.2857	0.189458
f_{n2}	"	1.70541	1.14449
f_{n3}	"	10.3125	6.8775
f_{n4}	"	16.5029	9.91715
f_{n5}	"	35.39	23.2937

TABLE 2 - D.F.T. process data specifications

Datum	Dim.	Value	Ref.
BW	rad/sec	393.672	Eq. (27)
NF	N.D.	32	" (22)
RF	rad/sec	12.30225	" (28)
T_s	sec.	0.5960	" (24)
S_t	sec.	0.510735	" (26)

Referring to the preceding data specifications the following processor characteristics are defined in order to evaluate the computational effort involved in the F.F.T. process.

- Data word length (NB) - The digital data word length determines the accuracy in the F.F.T. computation, i.e. the maximum error that can be found in the output spectrum numerical values. The data word length is chosen to satisfy the basic requirement established for the maximum allowable error for the observed output variable and incremented by an additional number of digits in order to compensate for the lost in numerical precision exhibited in the successive F.F.T. computational stages. Indicating with E the maximum error predicted for the observed state variable, the word length required for the accuracy will be given by:

$$E = \frac{1}{2^{NB} - 1} \quad (29)$$

The size word growth necessary to maintain the desired accuracy through the F.F.T. process can be evaluated in one bit for every two F.F.T. computational stages.

- Number of operations per butterfly (N_{OB}) - For a general complex input forcing function each butterfly requires four real multiplications and 2 real additions
- Number of machine cycle per operation (C_O) - For the class of microprocessor considered in this study, the sum of two successive products takes two machine cycle (T_C) and the data handling in and out the working memory locations, as required to apply the "in place" F.F.T. technique, is performed in four machine cycles.
- Number of machine cycles per butterfly (C_B) - For the considered microprocessor, six machine cycles are required per a butterfly computation
- Number of butterfly blocks per F.F.T. computation (N_B) - It is expressed by:

$$N_B = \frac{M}{2} \log_2 M \quad (30)$$

- Number of machine cycles for F.F.T. computation (C_F):

$$C_F = N_B \cdot C_B \quad (31)$$

- Time required for F.F.T. computation (T_F):

$$T_F = C_F \cdot T_C \quad (32)$$

- Static storage (S_{ROM}) - This requirement depends on the number of coefficients involved in F.F.T. and digital filtering algorithms to be stored in the processor R.O.M. memory area
- Temporary storage (S_{RAM}) - The R.A.M. area is dimensioned in function of the total number of words involved in the transit data flow.

Referring to the above mentioned definitions, the general characteristic of the F.F.T. dedicated microprocessor considered in this study are shown in Table 3.

TABLE 3 - F.F.T. processor characteristics

Datum	Dim.	Value	Ref.
NB	N.D.	8	Eq. (29)
S_{ROM}	K.bytes	1	-
S_{RAM}	K.bytes	2	-
M	N.D.	32	Eq. (22)
T_C	10^{-6} sec.	1	-
N_{OB}	N.D.	6	-
C_B	N.D.	6	-
N_B	N.D.	160	Eq. (30)
T_F	10^{-3} sec.	0,96	Eq. (32)
T_{tot}	10^{-3} sec.	3,2	-

In Table 3 is also indicated the total time (T_{tot}) required to compute the filtered discrete feedback function $f(n)$ given in (14) from the analog data observed in a time slot S_t established in Table 2.

The M.C.U. servosystem design includes the following parameters:

- Electro-optic sensor resolution (R_s)
- Force servoactuator resolution (R_F)
- M.C.U. control stiffness (K_{ST})
- Sensor Transducer gain (K_T)
- ADC full scale voltage (FSV)
- ADC resolution (R_{ADC})
- Servoactuator gain (K_i)

The forces applied to the M.C.U.s structure obey a sinusoidal distribution with an angular frequency fixed at 4.5 rad/sec.; These forces have the same phase but different magnitude depending on the M.C.U. location along the mast. To operate with different control authority at various grid points, the gains on the excitation signals, routed from the processor output to the M.C.U.s, are regulated as indicated in Table 4

TABLE 4 - Control gains

Gain	Dim.	Value
R_s	mrad	9.68
R_F	N.	0.01
K_{ST}	N/rad	10^3
K_T	volts/rad	60
FSV	volts	10
R_{ADC}	volts	$39 \cdot 10^3$
K_i	rad/amp.	0.0837
K_p (M.C.U.-1)	N.D.	0.01144
K_p (M.C.U.-2)	N.D.	0.03432
K_p (M.C.U.-3)	N.D.	0.0572
K_p (M.C.U.-4)	N.D.	0.08151
K_p (M.C.U.-5)	N.D.	0.10295

6 - The modal control system simulation

In order to evaluate the control performances of the proposed modal control system, a simulation software program was developed; it consists of three major blocks reflecting the effective system configuration and characteristics. In the first block the structural model describing the elastic behaviour of the flexible mast in terms of the normalized coordinates is formulated. The program, in this stage, solves the eigenvalue problem and computes the unforced and forced dynamical time responses, the last one referred to a specified distributions of the active control forces, which are introduced as a set of input data directly delivered by the program section simulating the control process.

The forced response under an established forces distribution, is programmed by computing the normalized modal matrix consisting of orthonormal modal vectors, by means of which the generalized force vector be is defined. The uncoupled second order differential equations describing the system forced dynamics, are solved for the transformed generalized coordinates; after the transformation back to the original generalized coordinates, the complete response solution is obtained and used to determine the linear and angular displacements of the mast reference section, which are the final output data for the structural program block. The second program block is devoted to solve the F.F.T. algorithm yielding the discrete Fourier transform of the feedback function which, when filtered, transformed in time domain and weighed by appropriate gains, yields the control levels used as input for the next program block; here the functional characteristics of the active control process, as it takes place in the central processor, are simulated. The force distributions in module and sign at the different mast grid points obtained in this program section are used as input data to the structural program block computing the mast's forced responses. From the program output data the control system performances in terms of the antenna pointing accuracy, can be evaluated by inspection. As reference objective in the design, the antenna line of sight pointing accuracy of the order

of 142 arc-second, as expected in the L.S.S. space mission to provide adequate communication service from a low earth orbit (500 Km.) has been assumed.

7 - Simulation results and conclusions

In Table 5 the simulation results, expressed in terms of antenna angular displacements averaged in the time slot established for the spectral process, are shown. Two cases are considered:

- Averaged antenna angular displacement with the active modal control inoperative:
 $(\alpha_r)_{INOP}$
- Averaged antenna angular displacement with the active modal control operative with the gains regulated as indicated in the preceding section: $(\alpha_r)_{OP}$

In both cases the initial conditions were fixed considering the mast deflected from its reference position to assume a static deformation obtainable with a unit force applied to the mast tip.

The values given in Table 5 are relative to the simulation referred to the first bending mode of vibration. In the same Table is indicated the averaged angular displacement $(\alpha_r)_M$ of the master section of the assumed model structure in its physical reality and submitted to the same initial conditions.

TABLE 5 - Antenna angular displacement
Averaged on $S_t = 0.510735$ sec.

$(\alpha_r)_{INOP}$	deg.	0.0147
$(\alpha_r)_{OP}$	"	0.0054193
$(\alpha_r)_M$	"	0.005956
E_{mf}	%	8.9
$\Delta\alpha_r$	%	36.8

In Table 5 are indicated also the percent error (E_{mf}) observed in the model following process and the percent antenna average angular ($\Delta\alpha_r$) displacement reduction obtained with the active modal control in respect to the value observed for the bare mast. The dynamical behaviour of the actual mast with reference to the model structure is presented in Fig. 7.

These results and others not presented here for reason of report length, indicate that the proposed active modal control is a feasible process capable to satisfy the requirement indicated for L.S.S. space mission.

REFERENCES

- 1 - A. Danesi - "A digital model following control system with discrete fourier transformed states" - Preprint VII Congresso Nazionale Associazione Italiana di Aeronautica ed astronautica, Napoli, 1983.
- 2 - Large Space Systems Technology - NASA Conference Publication 2215, November 1981.
- 3 - R.J. Karwosky - "An introduction to a Digital spectrum analysis including a high speed FFT processor design" TWR publication Note, 1980.
- 4 - Ali Zaheer - "High Speed FFT processor" - IEEE Transaction on Communications, May 1978.

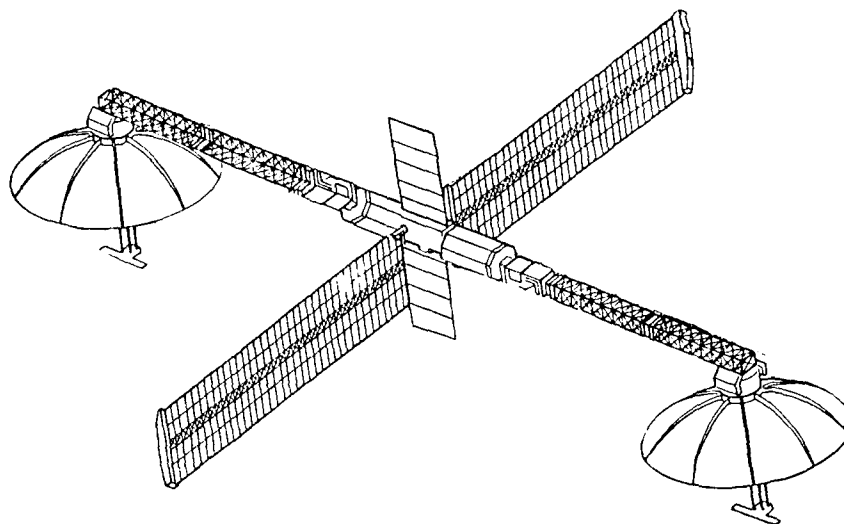


FIG. 1 L.S.S. Platform

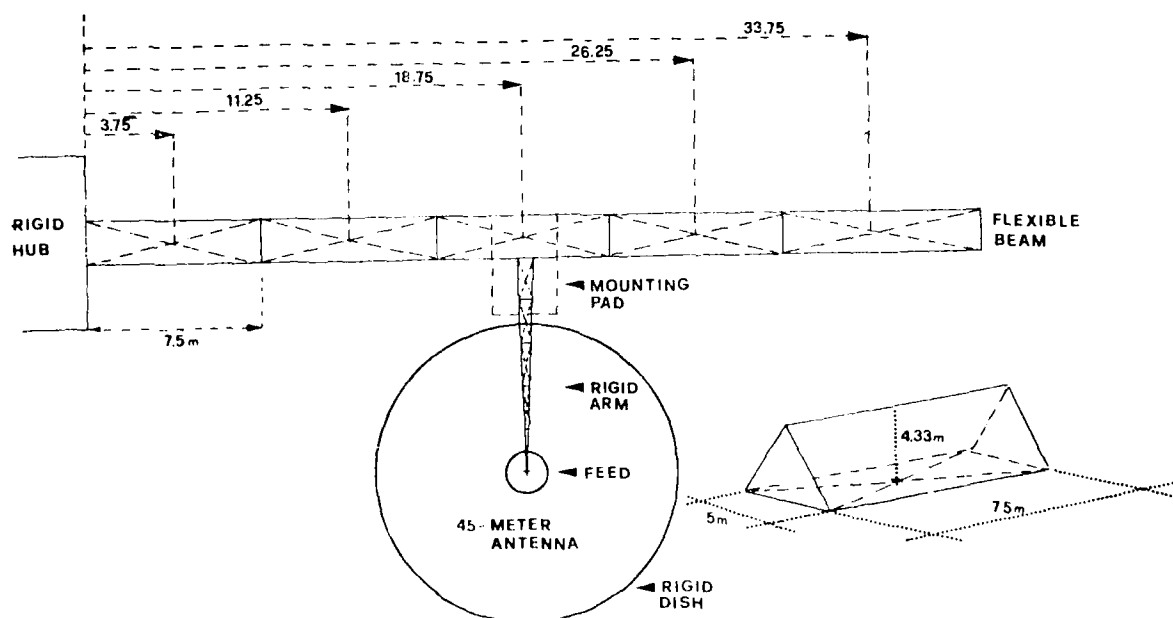


FIG. 2 5DOF Structural model

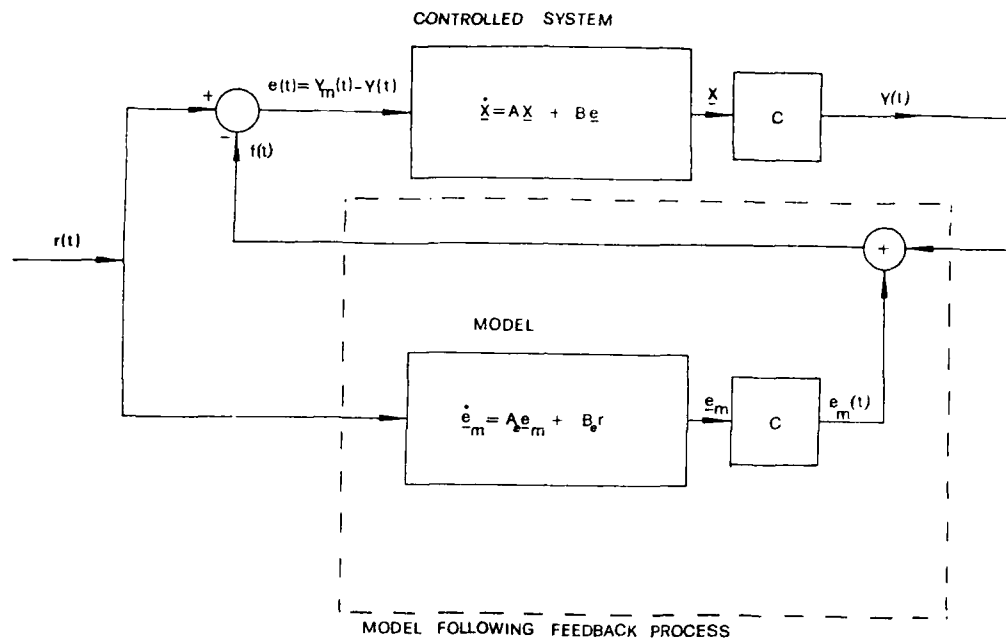


Fig 3 MODIFIED MODEL-FOLLOWING STRUCTURE

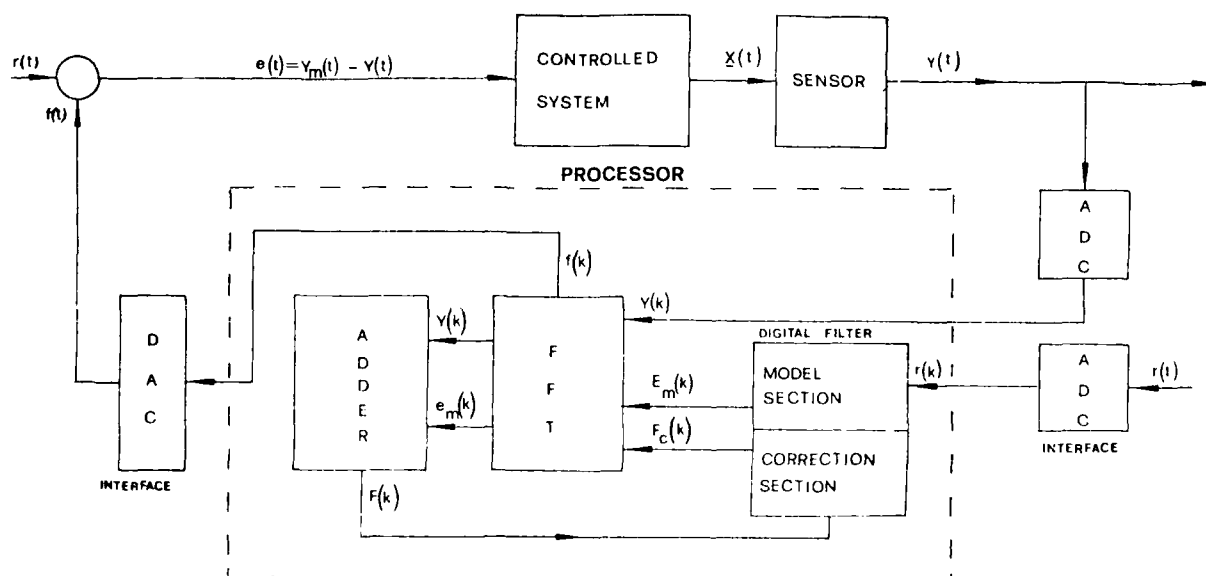


Fig 4 SPECTRAL PROCESSING

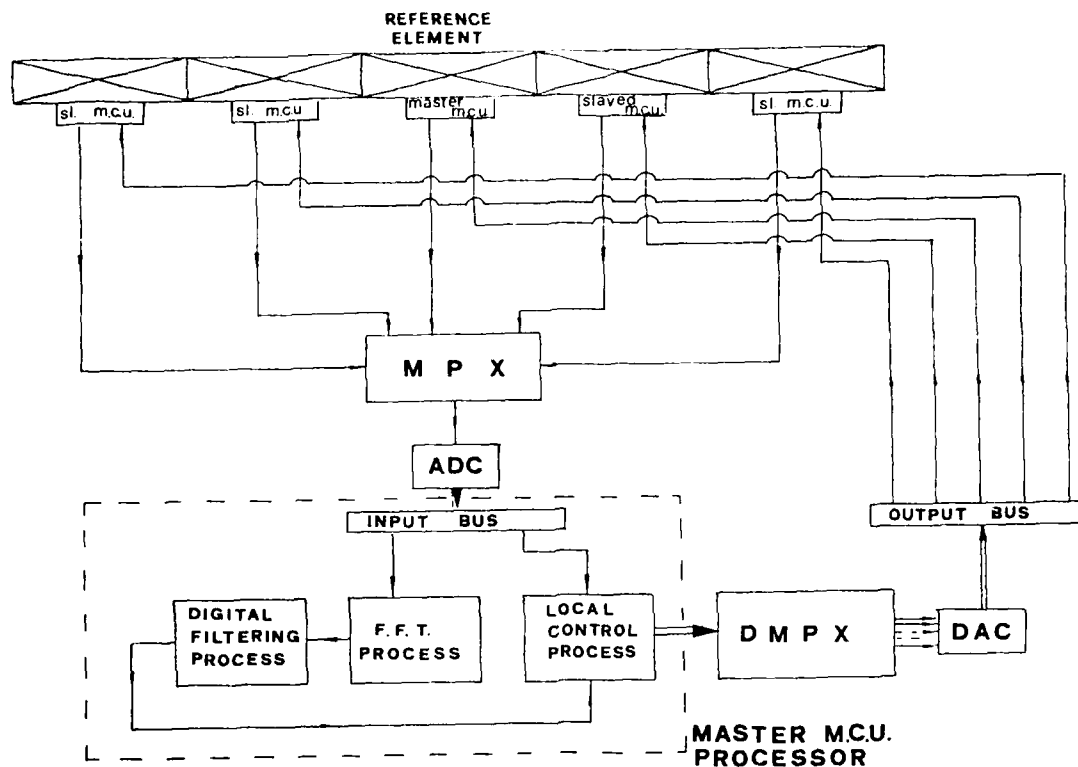


FIG. 5 Active modal control (A.M.C.) system

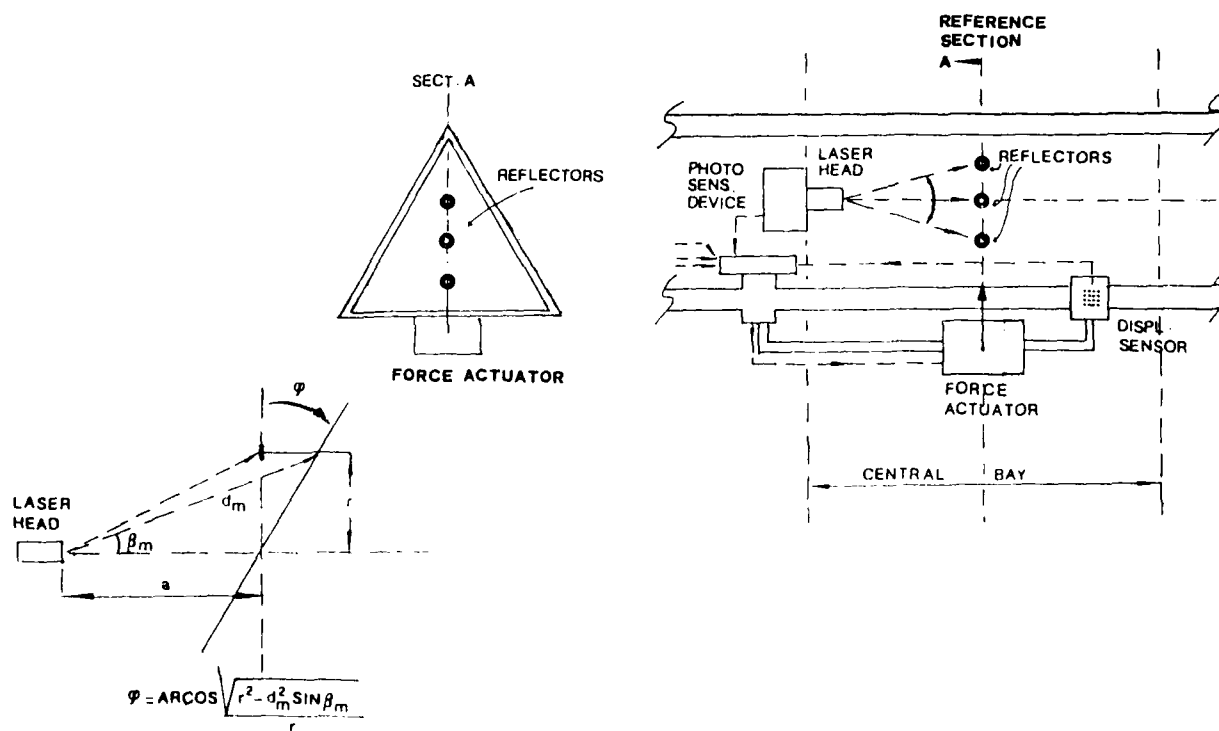
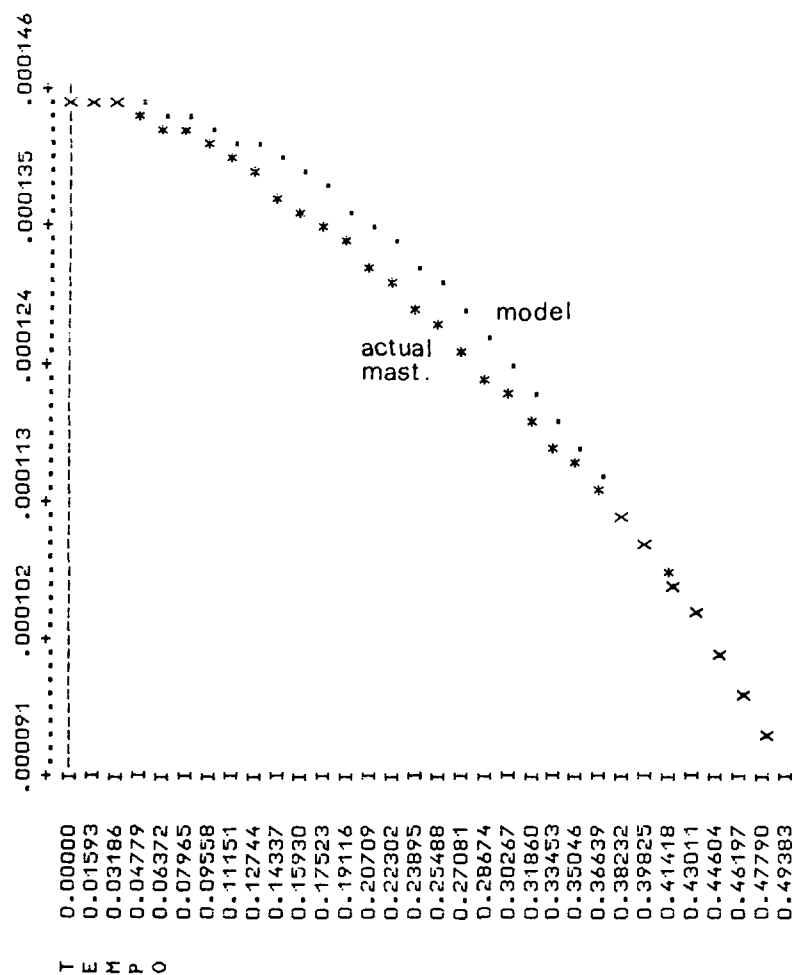


FIG. 6 A.M.C. System Implementation

FIG.7 Antenna ang. displ.
behaviour



T E M P O

AD P003 398

DESIGN AND ROBUSTNESS ANALYSIS OF REDUCED ORDER CONTROLLERS FOR LARGE FLEXIBLE SPACE VEHICLES

Siva S. Banda*, Lt D. Brett Ridgely*, Hsi-Han Yeh⁺, and Capt Douglas V. Palmer*

Flight Dynamics Laboratory, AFWAL/FIGC

Wright Patterson Air Force Base, Ohio 45433

U.S.A.

SUMMARY

In the control design of a large flexible space vehicle, a finite element model is truncated and the control system is designed on the basis of the reduced order model. This paper discusses the step - by - step application of frequency-shaped Linear - Quadratic - Gaussian methodology, as well as payoffs and costs of this method. The procedure for choosing and forming both state and control frequency - weightings is shown. Treating the unmodelled dynamics of the structure as a source of plant uncertainty, stability robustness evaluation was discussed. Practical usefulness of the singular-valued closed-loop performance analysis and its possible improvement are also discussed.

1. INTRODUCTION

Designing control systems for large scale systems, in particular, large space structures, has produced a great deal of discussion in the control community recently [1 - 5]. Some of the major problems that have been highlighted are 1) the difficulty of accurately modeling the structure, 2) trying to design control systems for the resulting very high order system model, usually resulting in the need to reduce the model order in order to design and implement the control system, 3) the fact that many of the resultant structural modes lie within the bandwidth of the controller, and will be affected by its use, 4) the extremely low natural damping of the structure, allowing any "misplaced" control energy (spillover) to easily drive the system unstable, 5) the fact that, in general, the more the controlled modes are forced into the left-half s-plane, the more "misplaced" control energy there is likely to exist, and 6) the fact that the structure must be autonomous, while sensors and actuators for this type of application are likely to be quite costly, thereby requiring careful consideration when determining the number of sensors and actuators to be used. There are, of course, other factors involved which require attention.

The paper is organized as follows. Section 2 gives the background for the reduced order control design and shows the need to shape the quadratic weights. Section 3 describes the frequency-shaped LQG design methodology. The effects of shaping the state weighting alone, control weighting alone, and both state and control weighting on the design are explained in Section 4. It also discusses the relationship between shaping the state/control weightings and the use of dynamic compensators/ shaping filters. Control systems are designed for two simple examples in Section 5. Section 6 discusses the use of singular values for robustness evaluation and for analyzing the closed-loop performance. Section 7 considers the first example from Section 5 and explains the difficulty in using the singular values for robustness evaluation. It also explains the need for integral control. Finally, Section 8 summarizes the conclusions of this paper.

2. BACKGROUND

In this paper, attention will be given solely to design methods which use finite element models of the structure versus those which start with partial differential equation representations. Obviously, since the actual structure has an infinite number of modes, the initial model, though very high order (200 or more modes for extremely large structures), still has modeling errors. These are errors which cannot be modeled and thus require a very robust control system if they are, in fact, important to the stability and performance of the structure. These errors will not be expressly addressed here, but would need to be considered before any real implementation of a control system.

Assuming now that a very high-order finite element model of the structure is available, a controller must be designed. It is reasonable to assume that the high frequency end of the model is not very accurate, due to limited knowledge of the structural behavior at high frequencies. Due to computational considerations (software routines and actual processor hardware) and the high frequency inaccuracies, the large model must be reduced in order to allow control design and implementation. Several methods for doing this have been proposed [5 - 6] which take into account the "weight", controllability, observability, etc., of the various modes. This process results in a "design model", versus the previous large model, called the "evaluation model".

Now a model is available to perform a control design upon. Linear-Quadratic-Gaussian (LQG) methods appear to be well suited to the multivariable nature of the problem, and will form the basis for the method proposed here. However, due to the complex nature of the problem, there are problems associated with the use of standard LQG design. The evaluation model may be represented by state and output equations of the form

$$\begin{bmatrix} \dot{x}_p \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_p & 0 \\ 0 & A_r \end{bmatrix} \begin{bmatrix} x_p \\ x_r \end{bmatrix} + \begin{bmatrix} B_p \\ B_r \end{bmatrix} u + \begin{bmatrix} D_p \\ D_r \end{bmatrix} w \quad (1)$$

* Aerospace Engineer

+ Associate Professor, on leave from University of Kentucky

$$y = [C_p \quad C_r] \begin{bmatrix} x_p \\ x_r \end{bmatrix} + v \quad (2)$$

where

- x_p - primary state vector
- x_r - residual state vector
- u - control vector
- w - process noise vector
- v - measurement noise vector

Assume now that the design model is of the form

$$\dot{x}_p = A_p x_p + B_p u + D_p w \quad (3)$$

with measurements

$$y_p = C_p x_p + v \quad (4)$$

The control u is selected to minimize the performance index

$$J = \int_0^\infty (x_p^T Q_1 x_p + u^T R_1 u) dt \quad (5)$$

and is given by

$$u = -K \hat{x}_p \quad (6)$$

where

$$K = R_1^{-1} B_p^T E_1 \quad (7)$$

$$0 = A_p^T E_1 + E_1 A_p - E_1 B_p R_1^{-1} B_p^T E_1 + Q_1 \quad (8)$$

with Q_1 and R_1 being constant, symmetric, positive semi-definite and positive definite matrices, respectively, chosen by the designer [7]. \hat{x}_p is an estimate of the controlled state produced through a Kalman filter, given by

$$\dot{\hat{x}}_p = A_p \hat{x}_p + B_p u + M [y - \hat{y}] \quad (9)$$

$$y = C_p \hat{x}_p \quad (10)$$

where

$$M = E_2 C_p^T R_2^{-1} \quad (11)$$

$$0 = E_2 A_p^T + A_p E_2 - E_2 C_p^T R_2^{-1} C_p E_2 + D_p Q_2 D_p^T \quad (12)$$

with Q_2 and R_2 being the strengths of the zero-mean, white, Gaussian process and measurement noises, respectively. Substituting the control law and augmenting the state estimates produces the closed-loop design model

$$\begin{bmatrix} \dot{\hat{x}}_p \\ \dot{\hat{x}}_r \\ \dot{\hat{x}}_s \end{bmatrix} = \begin{bmatrix} A_p & -B_p K \\ MC_p & A_p - B_p K - MC_p \end{bmatrix} \begin{bmatrix} \hat{x}_p \\ \hat{x}_r \end{bmatrix} + \begin{bmatrix} D_p & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} \quad (13)$$

which can be adjusted for closed-loop stability/performance through the choice of Q_1 and R_1 . However, this control law must be substituted into the evaluation model, as it is a more accurate model of the real structure. Doing this yields

$$\begin{bmatrix} \dot{\hat{x}}_p \\ \dot{\hat{x}}_r \\ \dot{\hat{x}}_s \end{bmatrix} = \begin{bmatrix} A_p & -B_p K & 0 \\ MC_p & A_p - B_p K - MC_p & MC_r \\ 0 & -B_r K & A_r \end{bmatrix} \begin{bmatrix} \hat{x}_p \\ \hat{x}_r \\ \hat{x}_s \end{bmatrix} + \begin{bmatrix} D_p & 0 \\ 0 & M \\ D_r & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} \quad (14)$$

The terms $B_r K$ and MC_r are called control and observation spillover, respectively. As can be easily seen, these terms affect the location of the poles of the design model (equation [13]) as well as the suppressed poles (eigenvalues of A_r), and they affect the performance as well. Since the suppressed poles have very

low open-loop damping and are high frequency, they lie very close to the imaginary axis in the s-plane. If they are "pushed" to the right by the spillover at all, they quickly go unstable. Readjusting Q_1 and R_1 to prevent this (the only means available for the given design model) can cause a drastic loss of closed-loop damping in the design model poles, in some cases to the point where almost no closed-loop improvement in damping is possible.

Note that as long as a Kalman filter (or an observer) is used, both control and observation spillover are present. When all of the states are available for measurement, only control spillover will occur. In this case, the closed-loop stability will not be affected by the suppressed modes, but the performance will. In other words, the closed-loop evaluation model's eigenvalues will be those of the design model plus those of the residual model by themselves, but the resulting eigenvectors will be different, thus affecting performance.

There are a number of ways to overcome the stability problem when a Kalman filter is present. By properly selecting K or by properly selecting actuator locations, the rows of the B_r matrix can be made orthogonal to the columns of the K matrix [8]. Another method is to minimize the control spillover by including the term $B_r u$ in the cost function J which is being minimized [3]. The problem, however, is arising from the fact that the penalty matrices Q_1 and R_1 penalize the states and controls by the same amount at all frequencies. A method to overcome this problem is now presented.

3. FREQUENCY-SHAPED LQG DESIGN

One way to avoid constant penalties is the use of frequency-shaped cost functionals [7], which is an extension of standard LQG design. In this method, the performance index to be minimized is assumed to be a function of frequency as follows

$$J = \frac{1}{2\pi} \int_{-\infty}^{\infty} [x_p^* (j\omega) Q_1(j\omega) x_p(j\omega) + u^*(j\omega) R_1(j\omega) u(j\omega)] d\omega \quad (15)$$

Note that here the weighting matrices Q_1 and R_1 are functions of frequency, rather than constant matrices. Furthermore, for ease of application, they are assumed to be rational functions of frequency squared, ω^2 . $Q_1(j\omega)$ must be positive semidefinite and $R_1(j\omega)$ must be positive definite. It is relatively simple (see Appendix A) to do a spectral factorization on Q_1 and R_1 to produce

$$Q_1(j\omega) = P_1^*(j\omega) P_1(j\omega) \quad (16)$$

$$R_1(j\omega) = P_2^*(j\omega) P_2(j\omega) \quad (17)$$

where P_1 and P_2 are rational matrices containing all of the resulting left-half plane poles and zeros. Both P_1 and P_2 , in general, have no restrictions upon the number of poles and zeros they contain. Now defining

$$P_1(j\omega) x_p = x^l \quad (18)$$

$$P_2(j\omega) u = u^l \quad (19)$$

the performance index of (15) may be written [using equations (16 and 17)] as

$$J = \frac{1}{2\pi} \int_{-\infty}^{\infty} [x^{l*}(j\omega) x^l(j\omega) + u^{l*}(j\omega) u^l(j\omega)] d\omega \quad (20)$$

P_1 and P_2 may be thought of as transfer functions to shape the old state and control vectors, x_p and u , into new ones, x^l and u^l . Therefore, minimal realizations of these transfer functions may be obtained.

These realizations take on slightly different forms depending upon the relative number of poles and zeros in the transfer functions. From equation (18), a general expression is given by

$$\dot{z}_1 = F_1 z_1 + G_1 x_p \quad (21)$$

$$x^l = H_1 z_1 + D_1 x_p + \sum_{i=1}^{j_1} D_{1i} x_p^{(i)} \quad (22)$$

where z_1 is a new additional state vector, j_1 is the number of zeros minus the number of poles in P_1 , and $x_p^{(i)}$ is the i th derivative of x_p . If j_1 is negative, $D_1 = 0$ and the summation is zero. If j_1 is zero (equal number of poles and zeros) the summation is zero and there is a feedforward term x_p . When j_1 equals one, the summation term contains a feedforward term \dot{x}_p , which can be replaced by the original design model

state equations, and this case may still be handled in a straightforward manner. Defining j_1 greater than one requires knowledge of the second or higher derivative of x_p and will not be developed here, as no practical applications have been found for this case.

The realization for equation (19) is slightly more complicated and requires further examination. If P_2 has a greater than or equal number of poles than zeros the development is straightforward and equation (19) may be represented by

$$\dot{z}_2 = F_2 z_2 + G_2 u \quad (23)$$

$$u^l = H_2 z_2 + D_2 u \quad (24)$$

If the chosen P_2 is of this category, and P_1 has no more than one excess zero, all possible choices of P_1 and P_2 result in developments of a similar form, and will be referred to as Form 1. If, however, P_2 has more zeros than poles, the form changes slightly and will be referred to as Form 2. The different developments will not be shown in detail in this paper; only the final results will be given.

3.1 FORM 1

The open-loop augmented system is

$$\begin{bmatrix} \dot{x}_p \\ \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} A_p & 0 & 0 \\ G_1 & F_1 & 0 \\ 0 & 0 & F_2 \end{bmatrix} \begin{bmatrix} x_p \\ z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} B_p \\ 0 \\ G_2 \end{bmatrix} u + \begin{bmatrix} D_p \\ 0 \\ 0 \end{bmatrix} w \quad (25)$$

$$y = [C_p \quad 0 \quad 0] \begin{bmatrix} x_p \\ z_1 \\ z_2 \end{bmatrix}^T + v \quad (26)$$

or
$$\dot{\tilde{x}} = A^1 \tilde{x} + B^1 u + D^1 w \quad (27)$$

$$y = C^1 \tilde{x} + v \quad (28)$$

where A^1 , B^1 , D^1 , C^1 , and \tilde{x} are the corresponding matrices and vector in equations (25) and (26). The F_1 , G_1 , F_2 , and G_2 matrices are found from the realizations of the chosen weighting matrices (shown in detail in the examples). Since $j_1 = 1$ in this case, equations (22) and (24) may be rewritten as

$$x^l = [D_1 + D_{11}A_p \quad H_1 \quad 0 \quad D_{11}B_p] \begin{bmatrix} x_p \\ z_1 \\ z_2 \\ u \end{bmatrix}^T \quad (29)$$

$$u^l = [0 \quad 0 \quad H_2 \quad D_2] \begin{bmatrix} x_p \\ z_1 \\ z_2 \\ u \end{bmatrix}^T \quad (30)$$

Substituting (29) and (30) into (20) and transforming to the time domain yields the expressions for \tilde{Q}_1 , \tilde{R}_1 , and N

$$J = \int_0^\infty \begin{bmatrix} \tilde{x}^T & u^T \end{bmatrix} \begin{bmatrix} \tilde{Q}_1 & N \\ N^T & \tilde{R}_1 \end{bmatrix} \begin{bmatrix} \tilde{x} \\ u \end{bmatrix} dt \quad (31)$$

where

$$\tilde{Q}_1 = \begin{bmatrix} (D_1 + D_{11}A_p)^T (D_1 + D_{11}A_p) & (D_1 + D_{11}A_p)^T H_1 & 0 \\ H_1^T (D_1 + D_{11}A_p) & H_1^T H_1 & 0 \\ 0 & 0 & H_2^T H_2 \end{bmatrix} \quad (32a)$$

$$\tilde{R}_1 = (D_{11}B_p)^T (D_{11}B_p) + D_2^T D_2 \quad (32b)$$

$$N = \begin{bmatrix} (D_1 + D_{11}A_p)^T (D_{11}B_p) \\ H_1^T (D_{11}B_p) \\ H_2^T D_2 \end{bmatrix} \quad (32c)$$

The feedback law that minimizes equation (31) is determined as follows:

$$u = -K_1 \hat{x}_p - K_2 z_1 - K_3 z_2 \quad (33)$$

$$[K_1 \ K_2 \ K_3] = \hat{R}_1^{-1} [B^1 T_{E_1} + N^T] \quad (34)$$

$$0 = E_1 A^1 + A^1 T_{E_1} + \hat{Q}_1 - (E_1 B^1 + N) \hat{R}_1^{-1} (E_1 B^1 + N)^T \quad (35)$$

Note that equations (34) and (35) contain an N matrix as well as the usual Q_1 and R_1 . This is due to the cross-coupling between the \hat{x} vector and the u vector. The modified Riccati equation is transformed into one of a suitable form for computer solution in Appendix B.

The state estimation equations are

$$\dot{\hat{x}}_p = A_p \hat{x}_p + B_p u + M [y - \hat{y}_p] \quad (36)$$

$$\hat{y}_p = C_p \hat{x}_p \quad (37)$$

where

$$M = E_2 C_p^T R_2^{-1} \quad (38)$$

$$0 = E_2 A_p^T + A_p E_2 - E_2 C_p^T R_2^{-1} C_p E_2 + D_p Q_2 D_p^T \quad (39)$$

and Q_2 and R_2 are as in equations (12). The closed-loop evaluation model then becomes

$$\begin{bmatrix} \dot{x}_p \\ \dot{\hat{x}}_p \\ \dot{z}_1 \\ \dot{z}_2 \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_p & -B_p K_1 & -B_p K_2 & -B_p K_3 & 0 \\ MC_p & A_p - B_p K_1 - MC_p & -B_p K_2 & -B_p K_3 & MC_r \\ 0 & G_1 & F_1 & 0 & 0 \\ 0 & -G_2 K_1 & -G_2 K_2 & (F_2 - G_2 K_3) & 0 \\ 0 & -B_r K_1 & -B_r K_2 & -B_r K_3 & A_r \end{bmatrix} \begin{bmatrix} x_p \\ \hat{x}_p \\ z_1 \\ z_2 \\ x_r \end{bmatrix} + \begin{bmatrix} D_r & 0 \\ 0 & M \\ 0 & 0 \\ 0 & 0 \\ D_r & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} \quad (40)$$

Note that the third equation, when written out, is

$$\dot{z}_1 = F_1 z_1 + G_1 \hat{x}_p \quad (41)$$

This change (from x_p to \hat{x}_p) is necessary since it is assumed here that x_p is not available for measurement, and causes no problems in the control gain determination as it is standardly assumed that the two problems (control and estimation) are decoupled, i.e. $x_p = \hat{x}_p$. Also, note that both control and observation spillover are present -- however, the designer now may have more ability to reduce their affect by selecting different frequency shapings. This is best seen further on in the examples.

3.2 FORM 2

In this case the degree of the numerator of P_2 is higher than the degree of the denominator.

This requires feeding forward \dot{u} or a higher derivative of u . For a system with one more zero than pole, the frequency-shaping states for the control weighting are

$$\dot{z}_2 = F_2 z_2 + G_2 u \quad (42)$$

$$u^{\ell} = H_2 z_2 + D_2 u + D_{21} \dot{u} \quad (43)$$

which has the block diagram shown in Figure 1. The highlighted differentiator block is troublesome. An equivalent block diagram is shown in Figure 2. This avoids the problem by augmenting the open-loop equations with additional states

$$\dot{u} = u_o \quad (44)$$

Now the open-loop augmented system (including state weighting) becomes

$$\begin{bmatrix} \dot{x}_p \\ \dot{z}_1 \\ \dot{z}_2 \\ \dot{u} \end{bmatrix} = \begin{bmatrix} A_p & 0 & 0 & B_p \\ G_1 & F_1 & 0 & 0 \\ 0 & 0 & F_2 & G_2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_p \\ z_1 \\ z_2 \\ u \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_o + \begin{bmatrix} D_p \\ 0 \\ 0 \\ 0 \end{bmatrix} w \quad (45)$$

$$y = [C_p \ 0 \ 0 \ 0] [x_p \ z_1 \ z_2 \ u]^T + v \quad (46)$$

which has the same form as equations (27) and (28), with u replaced by u_o . Assuming the state weighting has an equal or greater number of poles than zeros, the new state and control vectors become

$$x^\lambda = [D_1 \ H_1 \ 0 \ 0 \ 0] [x_p \ z_1 \ z_2 \ u \ u_o]^T \quad (47)$$

$$u^\lambda = [0 \ 0 \ H_2 \ D_2 \ D_{21}] [x_p \ z_1 \ z_2 \ u \ u_o]^T \quad (48)$$

The performance index has the same form as equation (31), with u replaced by u_o . The corresponding weighting matrices are

$$\begin{bmatrix} \tilde{Q}_1 & N \\ N^T & \tilde{R}_1 \end{bmatrix} = \begin{bmatrix} D_1^T D_1 & D_1^T H_1 & 0 & 0 & 0 \\ H_1^T D_1 & H_1^T H_1 & 0 & 0 & 0 \\ 0 & 0 & H_2^T H_2 & H_2^T D_2 & H_2^T D_{21} \\ 0 & 0 & D_2^T H_2 & D_2^T D_2 & D_2^T D_{21} \\ 0 & 0 & D_{21}^T H_2 & D_{21}^T D_2 & D_{21}^T D_{21} \end{bmatrix} \quad (49)$$

The feedback law which minimizes the performance index is given by

$$u_o = -K_1 x_p - K_2 z_1 - K_3 z_2 - K_4 u \quad (50)$$

where

$$[K_1 \ K_2 \ K_3 \ K_4] = \tilde{R}_1^{-1} [B^T E_1 + N^T] \quad (51)$$

and the necessary Riccati equation has the same form as equation (35). The Kalman filter equations have the same form as in equations (36) - (39). Now the closed-loop evaluation model can be formed, and becomes

$$\begin{bmatrix} \dot{x}_p \\ \dot{\hat{x}}_p \\ \dot{z}_1 \\ \dot{z}_2 \\ \dot{u} \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_p & 0 & 0 & 0 & B_p \\ MC_p & A_p - MC_p & 0 & 0 & B_p \\ 0 & G_1 & F_1 & 0 & 0 \\ 0 & 0 & 0 & F_2 & G_2 \\ 0 & -K_1 & -K_2 & -K_3 & -K_4 \\ 0 & 0 & 0 & 0 & B_r \end{bmatrix} \begin{bmatrix} x_p \\ \hat{x}_p \\ z_1 \\ z_2 \\ u \\ x_r \end{bmatrix} + \begin{bmatrix} 0 \\ MC_s \\ 0 \\ 0 \\ 0 \\ A_r \end{bmatrix} \begin{bmatrix} x_p \\ \hat{x}_p \\ z_1 \\ z_2 \\ u \\ x_r \end{bmatrix} + \begin{bmatrix} D_p & 0 \\ 0 & M \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ D_r & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} \quad (52)$$

While the control and observation spillover cannot be directly affected by choice of the weightings, the closed-loop design model is affected and can be made less sensitive to the spillover effects.

The two forms presented represent the most complex versions of each that a designer would practically want to choose. These were presented because any less complicated choice of weights results in a very simple change in the overall results. For example, if the control weighting meets the requirements of Form I, but there is no x_p feedforward term, D_{11} simply becomes zero. Further, if there is no x_p feedforward term, D_1 also becomes zero. Going one step further, if the control weighting is a constant matrix, the z_2 states and their associated terms in the penalty matrix are removed, and \tilde{R}_1 is simply the chosen constant. Finally if the state weighting is also constant, the z_1 states and corresponding terms

are removed, and the problem reduces completely to that of equations (3) - (14). This simplification also applies to weightings of Form 2. Note, however, that Form 2 does not simply reduce to Form 1 when a proper control weighting is used.

4. ANALYSIS OF THE METHOD

In this section, the effects of choosing different types of state and control weightings will be examined. First, state weighting alone will be examined, then control weighting alone, and then a combination of both. Tie-ins to classical type design techniques will be made throughout to give the reader a better understanding of the weightings.

4.1 STATE WEIGHTING ALONE

The block diagram for state weighting is shown in Figure 3. For the sake of clarity, the Kalman filter is left out. From this figure it is easily seen that

$$u = -K_1 x - K_2 z_1 \quad (53)$$

and

$$\dot{z}_1 = F_1 z_1 + G_1 x \quad (54)$$

Taking the Laplace transform of both equations, solving (54) for z_1 in terms of x , and then substituting into (53) yields

$$u = -[K_1 + K_2(sI - F_1)^{-1}G_1]x \quad (55)$$

Equation (55) shows the form of new feedback. Since $P_1(j\omega)$ is chosen to be diagonal, F_1 is always a diagonal matrix. G_1 is a matrix that contains blocks of zeros and blocks of diagonal type matrices, so that the product $(sI - F_1)^{-1}G_1$ never has cross-coupled terms (i.e., no terms with different denominators combine). Therefore, even though K_2 is generally dense, there are still no cross-coupled terms. This makes an element-by-element analysis simple, and the analysis may be performed in a single-input single-output setting.

Let us look at the problem from a selection of weighting viewpoint.

Point #1 - If a mode is given a zero weighting, you are telling it not to be regulated at any frequency.

Point #2 - If a mode is given a constant weighting, you are telling it to be equally regulated at all frequencies, including frequencies at which its characteristics may not be well known. This type and the above both represent no frequency weighting.

Point #3 - If a mode is given a weighting that has more poles than zeros, it is a decreasing function of frequency, and therefore the mode will be "heavily" regulated at low frequencies, and have decreasing regulation as frequency increases, to the point where it has no regulation at all.

Point #4 - If a mode is given a weighting with equal order numerator and denominator, of the form

$$c \frac{(j\omega + a_1)(j\omega + a_2)\dots}{(j\omega + b_1)(j\omega + b_2)\dots} \text{ then at low frequency the weighting has the value } \frac{ca_1 a_2 \dots}{b_1 b_2 \dots} \text{ while at high frequency it has the value } c.$$

Depending on the a and b values this can represent an increase or decrease in desired regulation.

Point #5 - If a mode is given a weighting with more zeros than poles, then the amount of regulation will increase with frequency. This can be arranged in such a way that the regulation will tend toward infinity at finite frequencies also (notch filter).

In light of the above comments, what type of feedback do we get if we make some of these choices?

Type #1 - No frequency weighting gives constant gain feedback.

Type #2 - More poles than zeros in the weighting gives a lag or lead filter in the feedback. dependent upon the gains; one obvious form is proportional plus integral feedback.

Type #3 - In the case of one zero divided by one pole, the weighting is factored into a constant plus a strictly proper transfer function. The zero does not appear directly in the feedback, and the weighting appears as a lag or lead dependent upon the gains, while the constant terms affect Q_1 to alter the proportional feedback.

Type #4 - In the case of more zeros than poles, the zeros do not appear in the feedback. The improper transfer function is factored into the sum of an n th order polynomial and a strictly proper transfer function. This proper transfer function "appears" in the feedback, as in the previous form. The polynomial affects the Q_1 , R_1 , and N matrices to alter the constant feedback.

Two items of interest arise from this analysis that may not be immediately obvious. First, care must be exercised before attempting an integral control type weighting. Since the weightings are realized by augmenting additional states to the plant, while the original system may have been fully controllable, the augmented system may not be. If it is not, then only poles that lie in the open left-half s -plane may be added. The $P_1(j\omega)$ realization eliminates right-half s -plane poles, but a pole at the origin can appear.

If this mode is uncontrollable, the system is unstabilizable and therefore does not have a Riccati solution.

The second item is the fact that explicit zero selection in the feedback is not possible. Also, a

pole is always present and therefore proportional plus derivative feedback ($K_1 + K_2s$) is not possible. In space structures, this type of feedback would not be desirable, so this is no major drawback.

4.2 CONTROL WEIGHTING ALONE

The block diagram for control weighting alone, again without a Kalman filter, is shown in Figure 4. From this figure it is seen that

$$u_o = \dot{u} = -[K_1x + K_3u] \quad (56)$$

Taking the Laplace transform and rearranging yields

$$u = -[sI + K_3]^{-1}K_1x \quad (57)$$

Since K_3 is computed as an optimal gain, there is no way to explicitly specify its value, hence there is no way to specify the pole location in equation (57). Notice that the block diagram is drawn for a special case of Form 2 from the previous (design) section. This is the case when the weighting has only zeros, no poles. Several points need to be made about choosing control weightings. First, if a weighting with more zeros than poles is chosen for one weighting, all weightings must be chosen to have the same number of excess zeros. If not, R_1 in equation (49) will not have full rank, which it must have.

Second, using a control weighting with more poles than zeros is not possible unless the state weighting has more zeros than poles. If the state weighting does not have more zeros, then the D_{11} term in equations (32a,b,c) will be zero, and the control weighting with more poles will have $D_2 = 0$. Therefore, R_1 will again be singular. Third, control weightings with an equal number of poles and zeros (or constants) may be used, but care must be exercised when using these if all the controls are not weighted this way. From equation (32), the R_1 matrix is

$$(D_{11}B_p)^T(D_{11}B_p) + D_2^TD_2 \quad (58)$$

If all weightings have an equal order numerator and denominator, D_2 will be present for each control and R_1 will have full rank provided the first term if present, does not reduce the rank of $D_2^TD_2$. It may also be seen that if some weightings do not have equal orders, they must be combined with appropriate state weightings in such a way that equation (58) does have full rank.

The next and final result in this subsection is the most important. Figure 5 shows the block diagram for inclusion of a shaping filter in the forward path of a standard optimal control problem. Actuator dynamics can be thought of as having the form of shaping filters. By solving both of these problems using arbitrary matrices, it can be shown that the resulting control laws feeding into the plant (and therefore the resulting closed-loop systems) are identical if the individual control weightings are chosen as the reciprocals of the individual shaping filter transfer functions. That is, if it is desired to shape a control with a transfer function of the form

$$u = \frac{a}{j\omega + b} u_c \quad (59)$$

then choosing a weighting of the form

$$P_2(j\omega) = \frac{1}{a} (j\omega + b) \quad (60)$$

yields an identical result. Therefore, if a designer knows the form of a shaping filter he wishes to use, he knows the type of control weighting to yield the same result.

4.3 STATE AND CONTROL WEIGHTING

Figure 6 shows the block diagram for this case, with the Kalman filter included. The corresponding state and control portions are outlined for clarity. The only effect that the Kalman filter inclusion has upon the previous results is that the state weightings actually shape the state estimates, but the estimates hopefully resemble the original states (the point behind the filter), and the true states are assumed to be unavailable anyway. The results of the previous two subsections both apply in this case.

5 CONTROL DESIGN EXAMPLES

5.1 EXAMPLE 1

An extremely simple, non-physical system will be used to illustrate the use of the method. This example is not meant to be representative of any physical structure -- rather it will show the basic use of the method. This example is not meant to be representative of any physical structure -- rather it will show the basic use of the method on a model that has the general characteristics of a flexible space structure. The open-loop evaluation model, partitioned into its controlled and suppressed states, is

$$\begin{bmatrix} \dot{x}_{1p} \\ \dot{x}_{2p} \\ \dot{x}_{1r} \\ \dot{x}_{2r} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & -0.01 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -25 & -0.05 \end{bmatrix} \begin{bmatrix} x_{1p} \\ x_{2p} \\ x_{1r} \\ x_{2r} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} w \quad (61)$$

$$y = [0 \ 1; 0 \ -1] [x_{1p} \ x_{2p}; x_{1r} \ x_{2r}]^T + v \quad (62)$$

First note that the state distribution matrix (A) is block diagonal, which is a typical form for a finite-element modeled space structure. This model contains two modes: a lower-frequency mode with damping ratio $\zeta = 0.005$ and natural frequency $\omega_n = 1.0$, and a higher-frequency mode with $\zeta = 0.005$ and $\omega_n = 5.0$.

This frequency separation will be assumed to be sufficient to justify dividing the overall state vector into controlled and suppressed state vectors. The controlled states make up the design model.

The design model is given by the following state and output equations:

$$\begin{bmatrix} \dot{x}_{1p} \\ \dot{x}_{2p} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -0.01 \end{bmatrix} \begin{bmatrix} x_{1p} \\ x_{2p} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w \quad (63)$$

$$y_p = [0 \ 1] [x_{1p} \ x_{2p}] + v \quad (64)$$

Since there are 2 states and one input, five penalty function elements (four in the two-by-two Q_1 matrix and one in the scalar R_1) must be chosen. In this example, four different combinations of these chosen elements were examined. In all cases, the state weighting matrix (Q_1) was assumed to be diagonal (for standard optimal control, this causes no loss of generality [10] - it is assumed this applies to frequency weighting as well). Also, the first element in the state weighting matrix, which would represent a weight on the displacement of the mode, was chosen to be zero in all cases. Therefore, only two weights had to be chosen. The four cases were as follows:

- 1) Standard LQG - state and control weightings constant at all frequencies.
- 2) State Weighting - state weighting was made a function of frequency with control weighting constant.
- 3) Control Weighting - control weighting was made a function of frequency with state weighting constant.
- 4) State & Control Weighting - both state and control weightings were made functions of frequency.

In cases 2 and 4, all of the state frequency weightings chosen had the form

$$Q_1(j\omega) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{a^2}{\omega^2 b^2} \end{bmatrix} \quad (65)$$

and in Cases 3 and 4 all of the control frequency weightings had the form

$$R_1(j\omega) = c^2 (\omega^2 + d^2) \quad (66)$$

$Q_1(j\omega)$ was chosen as a decreasing function of frequency since the model is not well known at high frequency and therefore small penalties on state deviations at high frequency are desirable. Since the suppressed modes are not accounted for in the design model, the designer does not have any influence over how high-frequency control energy will affect those modes. Since they are very close to the imaginary axis ($-0.025 \pm 5j$, in this example), any control energy that causes even a small shift to the right in these poles may result in an unstable system. Therefore, the control weighting, $R_1(j\omega)$, was chosen to be an increasing function of frequency (where there are suppressed modes) will be heavily penalized.

The spectral factorizations of $Q_1(j\omega)$ and $R_1(j\omega)$ are almost trivial in these cases. It may be easily verified that

$$P_1(j\omega) = \frac{a}{j\omega + b} \quad (67)$$

$$P_2(j\omega) = c(j\omega + d) \quad (68)$$

The basic form of the four cases is shown in the first row of Table 1. The designer is free to choose any positive values for a, b, c, and d. For generality, these are left as variables in the following development. The control frequency weightings are of Form 2, so that development must be followed for Cases 3 and 4.

The next step, once $Q_1(j\omega)$ and $R_1(j\omega)$ are chosen and factored, is to determine realizations for the $P_1(j\omega)$ and $P_2(j\omega)$. The next row in Table 1 shows these results. For Cases 1 and 3 where there is no state frequency weighting, no z_1 states are required, and therefore there are no F_1 , G_1 , or H_1 matrices [from equations (21) and (22)] and D_1 is found from

$$Q_1 = D_1^* D_1 \quad (46)$$

A similar argument can be made for the control weightings. Since none of the weightings contain poles, no F_2 , G_2 , or H_2 matrices are needed as there are no z_2 states. For Cases 1 and 2, D_2 is simply the square root of R_1 . In Cases 3 and 4, both a D_2 and a D_{21} matrix (scalars in this case) are required to handle the weighting's zero.

The next row in Table 1 shows the resulting \tilde{Q}_1 , \tilde{R}_1 , and N matrices (Cases 1 and 2 do not need an N matrix). These are found from equation (31) [Cases 1 and 2] or (49) [Cases 3 and 4], after removing the rows and columns corresponding to vectors that are not used in the realizations. The fourth row shows the corresponding A^1 and B^1 matrices from equation (27), which are used in the appropriate Riccati equation to solve for the feedback gains. The form and dimension of these feedback gains are also shown.

The final row shows the closed-loop design A matrix, A_p , and its corresponding closed-loop evaluation matrix, A_E . These matrices are formed from reductions in equation (40) or (52), depending upon the form of the weightings.

The variables a, b, c, and d were tuned to define weightings that produced acceptable closed-loop eigenvalues. Only the closed-loop eigenvalues were examined for this model. Here two criteria were examined to determine acceptable results -- a closed-loop controlled damping ratio near 0.7; and a stable suppressed mode. Results for each case are shown in Table 2. These are not meant to be the absolute best choice of the weights; rather, they are representative of the type of achievable results. For all cases, the strengths of the noises w and v were chosen to be $\sigma_2 = 0.01$ and $\sigma_2 = 0.001$. This was a completely arbitrary selection.

Case 1, standard LQG, gave a good design damping ratio, but yielded an unstable suppressed mode. State frequency-weighting alone (Case 2) yielded the same result. Case 3, however, gave a stable evaluation model. Unfortunately, the design damping ratio could not be made as high as 0.7. The final case, with both state and control frequency weightings, gave the desired damping ratio and a stable suppressed mode. It is also interesting to note that the suppressed damping ratio was improved by a factor of over three from its open-loop value. This is not a goal here -- it is desired only to keep the system from becoming unstable. Other methods would then be used to improve the suppressed damping ratio, such as low authority control [5].

5.2 EXAMPLE 2

Now a slightly larger system will be used to illustrate application of the method. The system once again has the basic finite-element structure characteristic of a large space structure, given by

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -\omega_1^2 I & -2\zeta_1 \omega_1 I \end{bmatrix} x + \begin{bmatrix} 0 \\ B \end{bmatrix} u + \begin{bmatrix} \gamma \\ D \end{bmatrix} w \quad (70)$$

$$y = [0 \ C] x + v \quad (71)$$

where, for this example, i runs from one (1) to four (4), and the specific ω_1 values are

$$\omega_1 = (1.66, 2.96, 4.20, 9.25) \quad (72)$$

with $\zeta_1 = 0.005$ for all modes. The B, C, and D matrices were chosen arbitrarily to be

$$B = \begin{bmatrix} -6.90 \times 10^{-2} & -1.67 \times 10^{-2} \\ 2.49 \times 10^{-1} & -5.96 \times 10^{-2} \\ 2.89 \times 10^{-1} & 2.89 \times 10^{-1} \\ 3.17 \times 10^{-1} & -1.50 \times 10^{-1} \end{bmatrix} \quad (73)$$

$$C = \begin{bmatrix} 1.12 \times 10^{-1} & -1.89 \times 10^{-1} & 2.89 \times 10^{-2} & -1.67 \times 10^{-1} \\ 1.67 \times 10^{-2} & 2.49 \times 10^{-1} & 1.50 \times 10^{-1} & -2.89 \times 10^{-1} \end{bmatrix}$$

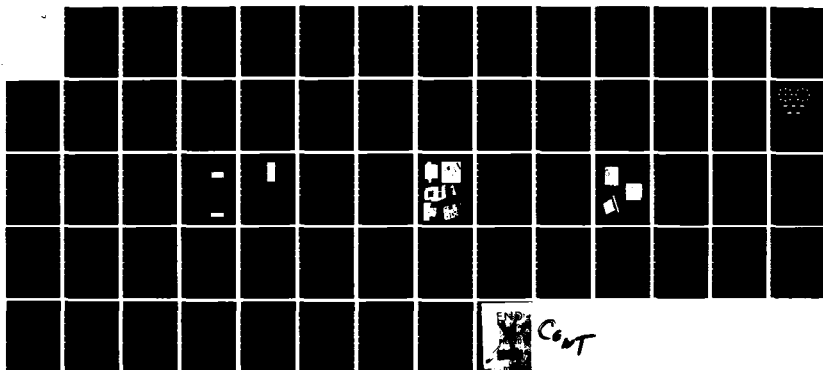
AD-A141 969

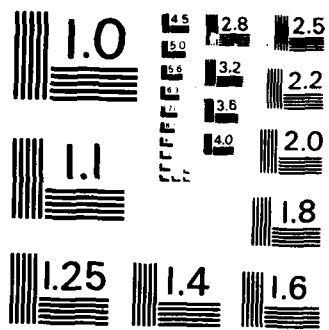
CONFERENCE PROCEEDINGS ON GUIDANCE AND CONTROL
TECHNIQUES FOR ADVANCED SP. (U) ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY. JAN 84
AGARD-CP-350 F/G 22/3

415

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (75)$$

By a completely arbitrary selection, the three lowest frequency modes were retained in the control design model and the highest frequency mode was suppressed. The design objective for this example was to yield a control law for which the closed-loop evaluation model had all primary modes with a damping ratio close to 0.2, and also had a stable suppressed mode. First, standard LQG design was used. For a choice of $Q_1 = \text{diag}\{0,0,0,10,10,10\}$, $R_1 = \text{diag}\{0.01, 1.0\}$, and with the process and measurement noises all having strengths equal to those in the previous example, the basic results are seen in Table 3, under the LQG design heading. Note that the highest frequency mode in the design model has been moved to the real axis and split apart. The most important fact is that the suppressed pole has gone unstable.

Next, several choices of state and control weightings were made. The final choice was

$$P_1 = \{\text{diag } 0.0, 0, \frac{50}{j\omega+4}, \frac{30}{j\omega+3}, \frac{25}{j\omega+2}\} \quad (76)$$

and $P_2 = \text{diag}\{0.2(j\omega+2), 0.2(j\omega+2.1)\}$. The new \hat{Q}_1 , \hat{R}_1 , and N matrices for this choice along with the basic results are given in the second section of Table 3. It is clearly seen here that all design poles have a damping ratio close to 0.2, and that the suppressed pole is stable (once again, with a higher damping ratio than open-loop). The gains for this case and for optimal control alone are on the same order of magnitude, with the maximum individual gains being on the order of 90 for optimal control and 70 for the frequency-weighted case. The additional poles due to the frequency weightings usually do not affect the closed-loop response as there are zeros very close to their locations. The overall performance of the system is discussed at length in Sections 6 and 7.

The rationale behind choosing the weightings was as follows. The lowest frequency mode is separated enough from the suppressed frequency so that the need for low frequency roll-off is not overwhelmingly critical, therefore a break frequency (pole) was chosen to be 2 radians per second. The intermediate design mode was given a pole in between these, at 3 radians per second. The zeros of the control weighting were chosen to be 2.0 and 2.1 radians per second to insure sufficient control roll-off. The numerators of all weightings were then tuned to achieve the desired results.

6. ROBUSTNESS EVALUATION

The uncertainty in the system modelling arises from several sources: (1) Parameter errors - these are present due to the imprecise knowledge of modes and mode shapes of the structure. (2) Model truncation errors - these are due to the unmodelled dynamics of the structure. (3) Nonlinearities - these may arise due to, for example, large angular rates of motion of the structure. (4) Disturbances - due to rotating machines, combustion devices, rotating antennas and other on-board equipment. In addition, gravity gradients and solar pressure are external disturbance sources. (5) Parameter errors and model truncation errors in the actuator/sensor dynamics. (6) Finally, the noise in the sensors and actuators may also contribute to the uncertainty in the system modelling. While it is important to design a control system such that it is robust in the presence of all these uncertainties, the only uncertainty that was studied in this paper was that due to the unmodelled dynamics. The objective of the rest of this paper is to evaluate the robustness of a reduced order control design method (in the presence of unmodelled dynamics) using singular value analysis.

There are numerous publications contributing to the theory of robustness of multivariable feedback systems [for example, references 11 and 12]. One can think of the actual structure as being a combination of the design model of the structure and uncertainty in the model. The purpose here is to evaluate the stability robustness and performance robustness (in the presence of the uncertainty) of the control system design using the method discussed in the previous section. The residual dynamics of the structure are modelled as the plant uncertainty in this paper. Hence a stability robustness test is required for evaluation.

Consider a multi-input multi-output control system as shown in figure 7. $G(s)$ is a precompensator transfer matrix, $H(s)$ is a feedback controller transfer matrix, $P(s)$ is the plant transfer matrix and $\Delta(s)$ is an uncertainty matrix. Assume dimensions of $G(s)$, $H(s)$, $P(s)$ and $\Delta(s)$ to be $(l \times m)$, $(m \times l)$, $(d \times l)$ and $(d \times l)$ respectively. Note that the uncertainty is included in Figure 7 as an additive perturbation. We can write the closed-loop transfer matrix $T.M.$, in three different forms:

$$T.M. = (P+\Delta) G [I+H(P+\Delta) G]^{-1} \quad (77)$$

$$= (P+\Delta) [I+GH(P+\Delta)]^{-1} G \quad (78)$$

$$= [I+ (P+\Delta) GH]^{-1} (P+\Delta) G \quad (79)$$

From the transfer matrices in equations (77) - (79), after making assumptions about the invertibility of the matrices involved, we can derive several stability robustness tests as

$$\bar{\sigma}(H\Delta G) < \underline{\sigma}[I + HPG] \quad (80)$$

$$\bar{\sigma}(H\Delta) < \underline{\sigma}[G^{-1} + HP] \quad (81)$$

$$\bar{\sigma}(\Delta G) < \underline{\sigma}[H^{-1} + PG] \quad (82)$$

$$\bar{\sigma}(\Delta) < \underline{\sigma}[(GH)^{-1} + P] \quad (83)$$

$$\bar{\sigma}(GH\Delta) < \underline{\sigma}[I + GHP] \quad (84)$$

$$\bar{\sigma}(\Delta GH) < \underline{\sigma}[I + PGH] \quad (85)$$

where $\bar{\sigma}[\cdot]$ and $\underline{\sigma}[\cdot]$ are the maximum and minimum singular value functions obtained by the singular value decomposition of the transfer matrices in the parenthesis. Equations (80) - (85) are all sufficient conditions for stability robustness of the closed-loop system when the residual dynamics are modelled as an additive perturbation.

In reference [4] Kosut derived a stability robustness test treating the residual dynamics as an input multiplicative perturbation to the plant. The residual dynamics of the plant can be treated as an output multiplicative perturbation as well. So, a different set of robustness tests can be derived treating the residual dynamics as either input multiplicative perturbation or output multiplicative perturbation. The stability robustness tests shown in equations (80) - (85) have been obtained using Doyle's formulation (reference 11). In a similar fashion, a whole new set of robustness tests can be derived using Lehtomaki's formulation (reference 12). The point is that all these robustness tests are sufficient conditions for robustness of the same system with the same uncertainty, the residual dynamics. One may ask the question - are these robustness tests equivalent? If the design fails one robustness test, will it fail the other? To demonstrate under what conditions one test implies another let us consider, for example, equations (83) and (85).

Let us suppose equation (85) holds. Then,

$$\bar{\sigma}(\Delta GH) < \underline{\sigma}[I + PGH] = \underline{\sigma}[(GH)^{-1} + P] GH \quad (86)$$

Using the inequality (C-1) from Appendix C, equation (86) implies

$$\bar{\sigma}(\Delta) \underline{\sigma}(GH) \leq \underline{\sigma}[(GH)^{-1} + P] \bar{\sigma}(GH) \quad (87)$$

or

$$\bar{\sigma}(\Delta) \leq \underline{\sigma}[(GH)^{-1} + P] \frac{\bar{\sigma}(GH)}{\underline{\sigma}(GH)} \quad (88)$$

Now let us suppose equation (83) holds. Then

$$\bar{\sigma}[\Delta GH(GH)^{-1}] < \underline{\sigma}[(I + PGH)(GH)^{-1}] \quad (89)$$

Again using the inequality (C-1) from Appendix C, equation (89) implies,

$$\bar{\sigma}(\Delta GH) \leq \underline{\sigma}(I + PGH) \frac{\bar{\sigma}(GH)}{\underline{\sigma}(GH)} \quad (90)$$

Thus from equations (88) and (90) we notice that a sufficient condition for equations (83) and (85) to be equivalent is that

$$\frac{\bar{\sigma}(GH)}{\underline{\sigma}(GH)} = 1 \quad (91)$$

Equation (91) holds if GH is a unitary matrix, which is generally not true for practical systems. In other words, for practical systems we can be certain that the stability robustness tests given in equations (83) and (85) are not equivalent. In a similar fashion, one can derive sufficient conditions for the other tests given in equations (80) - (85) to be equivalent and show that these tests do not imply each other for practical systems.

Since all the robustness tests are not equivalent, and since there is no way to tell which test is the best one for the uncertainty in question, we will only use the test from equation (80) after substituting H_p , H_r , and H_c in the place of P, Δ and H respectively. H_p and H_r indicate primary and residual transfer matrices of the structure. H_c is the controller transfer matrix to be designed using the method discussed in the previous section. We also assume that the pre-compensator G(s) is not present. Then we have

$$\bar{\sigma}(H_c H_r) < \underline{\sigma}[I + H_c H_p] \quad (92)$$

or

$$\frac{\underline{\sigma}[I + H_c H_p]}{\bar{\sigma}(H_c H_r)} > 1 \quad (93)$$

where, from equations (1) and (2) H_p and H_r are given by

$$H_p(s) = C_p [SI - A_p]^{-1} B_p \quad (94)$$

$$H_r(s) = C_r [SI - A_r]^{-1} B_r \quad (95)$$

The form of H_c depends on the selection of the frequency shaped quadratic weights in equation (15) and it is given by

$$H_c(s) = \beta(s) (sI - A_p - B_p \beta(s) + MC_p)^{-1} M \quad (96)$$

where

$$\beta(s) = -K_1 \quad (97)$$

or

$$\beta(s) = -[K_1 + K_2 (SI - F_1)^{-1} G_1] \quad (98)$$

or

$$\beta(s) = -(SI + K_4)^{-1} K_1 \quad (99)$$

or

$$\beta(s) = -(SI + K_4)^{-1} [K_1 + K_2 (SI - F_1)^{-1} G_1] \quad (100)$$

Equations (97) - (100) are for the cases when no frequency weighting is used, when state weighting alone is used, when control weighting (with no poles) alone is used, and when both state and control weightings are used, respectively. Note that in equations (97) - (100) the gain matrices K_1 , K_2 and K_4 are as defined in equation (59). It is clear that the choice of the frequency shaped state and control weightings affect the controller transfer function $H_c(s)$, which in turn affects the robustness test given in equation (93).

If the control system design passes the test in equation (93) it only guarantees stability robustness of the system. In the design process one likes to check the performance of the closed-loop system as well. For good tracking and (low frequency) disturbance rejection we would like to have $\sigma(H_c H_p)$ high at low frequencies. For minimizing the response to sensor noise, we would like $\sigma(H_c H_p)$ to be low (attenuation) at high frequencies. We also require performance tolerance to uncertainties, which are the residual dynamics. If $\sigma[H_c(H_p + \Delta)]$ where $\Delta = H_r$ is obtained in such a way that it will not deteriorate the performance as seen in $\sigma(H_c H_p)$ plot, then the closed-loop performance is said to be robust for the uncertainty in question. In other words, if the performance of the nominal system is close to the performance of the perturbed system, then performance robustness will be achieved. Of course, the performance of the nominal system must be acceptable or the robustness results are irrelevant.

7. EXAMPLE FOR ROBUSTNESS EVALUATION

Let us consider again the example shown in equations (61) and (62). The conclusions we draw from the discussion of this example hold, in general, for space structures with larger dimension. Control systems are designed using the method discussed in Section 3. An elaborate discussion of the approach for the control design for this example is given in Section 5. In this section we will only concentrate on the robustness issues of the design. Several control laws were designed using frequency shaping on state weighting alone, control weighting alone, or both. In all these designs, although the controller was designed for the primary modes, the closed-loop evaluation model (equation (52)) was stable. Thus it was decided to study the applicability of the stability robustness tests given in equation (80) - (85). The residual dynamics of the structure was modelled as the uncertainty. Because of the reasons explained in Section 6, the test given in equation (93) was selected as a stability robustness test. It was seen that each and every control system design failed the stability robustness test (except one, as mentioned below) although, in fact, the control system with the evaluation model was stable. As an example, Figure 8 shows the test when a state weighting of zero on the displacement and $100/(\omega^2+1)$ on the rate of displacement along with a control weighting of (ω^2+4) , was used.

It should be mentioned that the singular value test is only a sufficient condition meaning that if the design fails the test, the system really may or may not be stable. The point is that if one decides to iteratively select the quadratic weights to design a control law for the reduced order model and checks the robustness using singular value test, and if the design fails the test, the results may be erroneous. Figure 9 shows a sample of one design which passes this test. In this case a state weighting of zero on the displacement and $0.01/(\omega^2+1)$ on the rate of displacement only was used. Although this design passed the stability robustness test, its performance (as seen in figure 10) is quite unacceptable. The performance is unacceptable because this design resulted in gains so small that the closed-loop performance is almost the same as open-loop performance. Therefore, this case will not be considered when further conclusions are given in the rest of the paper.

The fact that the design failed the robustness test in all other cases can be explained as follows. The stability robustness tests shown in equations (80) - (85) assume that the uncertainty as measured by the singular value is in the worst possible direction. Since the uncertainty in the example problem was assumed to be known and since its direction (phase angle, because it is a single input-single output problem) was not in the worst direction equations (52) showed that it was a stable design where as the design failed the test in equation (93), indicating that the design may or may not be stable. It is our opinion that the singular value robustness tests as shown in equations (80) - (85) may be hard to use on

some systems (such as the one shown in equations (61) and (62)) from the applications point of view. If information about the direction and or structure of the uncertainty is somehow, included, then these tests may be less conservative and they could be useful in a design process.

However, a plot of the minimum singular value of the loop-gain transfer matrix is quite useful since it gives us a very quick estimate of the lower bound of the closed-loop performance. Numerical simulation of a very large set of differential equations that represent the closed-loop evaluation model (equation (52)) is a computational burden and does not have to be carried out for every quadratic weight selection - rather it can be carried out only when the (loop gain transfer matrix) is satisfactory. For multi-input multi-output systems (which large space structures always are), the closed-loop system will have several bandwidths - one for each loop. The cross-over frequencies of the minimum and maximum singular value plots of the loop gain will give an estimate of the range of all the bandwidths.

Although every control system designed for the example in equations (61) and (62) failed the robustness test, since we know several of those designs yielded stable evaluation models, performance of those designs was studied. Performance plots of the nominal and perturbed (with the uncertainty included) closed-loop systems are almost the same and have the same characteristics in all the designs. These plots typically look like the ones shown in figures 11 and 12 and correspond to the weighting selection made to obtain figure 8. Typically, all the $\underline{g}(H_c H_p)$ plots show a very small value (≈ 1) at low frequencies (≈ 0.02 rad/sec), indicating poor tracking and disturbance rejection properties. From figures 11 and 12 it can be easily seen that the roll-off frequency is very high (≈ 30 rad/sec) indicating a very poor performance at high frequencies. A careful examination as discussed below indicates the reasons for poor performance and suggests remedies for improving it.

When the model of the structure is represented as shown in equations (3) and (4), the plant transfer matrix $H_p(s)$ is as shown in equation (D-6) of Appendix D (Notice that the rigid body modes are ignored). Let $H_c(s)$ be represented by

$$H_c(s) = \frac{1}{\hat{d}(s)} \hat{N}(s) \quad (101)$$

where $\hat{d}(s)$ is a polynomial and is the least common multiple of the denominators of the individual elements of $H_p(s)$ and $\hat{N}(s)$ is an $m \times l$ matrix. From equations (D-6) and (101) the loop-gain transfer matrix is given by

$$H_c H_p = \frac{s}{d(s) \hat{d}(s)} \hat{N}(s) N(s) \quad (102)$$

It is the free s in the numerator of equation (102) that gives unacceptable performance at low frequencies. For good performance we need the denominator polynomial $\hat{d}(s)$ to have a free "s" to cancel the one in the numerator. Since $d(s)$ can never have a free "s" (see equations (D-5) and (D-7)), $\hat{d}(s)$ must have a free "s". In other words, the controller transfer function matrix (equation (101)) must have a free "s" in the denominator. That is, we need to use integral control. Therefore, for large structures that are represented by the state equations as shown in equation (3), regardless of what method we use to design a control system it is apparent that we always need to use integral control to obtain good performance of the closed-loop system at low frequencies as long as the measurement vector (equation (4)) contains only rates. In other words, if the displacements are measured instead of or along with the rates, integral control is not needed. This can be easily seen by rewriting equation (D-2) to measure displacements and rederiving equation (D-6). Figures 13 and 14 show the stability test and performance plot for an example with the state equations of equation (61) and an output equation as follows

$$y = [1 \ 0 \ -1 \ 0] [x_{1p} \ x_{2p} \ x_{1r} \ x_{2r}]^T + v \quad (103)$$

A state weighting of zero on the displacement and $49/\omega^2 + 4$ on the rate of displacement, along with a control weighting of $[\omega^2 + 1]$, was used. Figure 13 shows that the design failed the stability test, as expected. It can be seen, however, that good performance at low frequencies can be obtained using displacements as measurements.

Since in practice, though, we prefer to measure rates instead of displacements, let us now investigate if we can design integral control using frequency shaped cost functional methodology for a finite element model of a flexible structure with rates as measurements. From equations (96) and (99), it can be seen that it is difficult to use control weighting alone to obtain integral control. From equations (96) and (98) it seems easy to use state weighting alone to obtain integral control. To make the point, all we need to consider is the simplest example with one mode and one input. Then the matrices A_p and B_p in equation (3) are

$$A_p = \begin{bmatrix} 0 & 1 \\ -\omega_1^2 & -2\zeta_1 \omega_1 \end{bmatrix} \quad B_p = \begin{bmatrix} 0 \\ b_1 \end{bmatrix} \quad (104)$$

To obtain integral control, let us choose the state weighting as

$$Q(j\omega) = a/\omega^2 \quad (105)$$

Now the augmented matrices A^1 and B^1 in equation (27) are

$$A^1 = \begin{bmatrix} 0 & 1 & 0 \\ -\omega_1^2 & -2\zeta_1\omega_1^2 & 0 \\ 0 & a & 0 \end{bmatrix} \quad \text{and } B^1 = \begin{bmatrix} 0 \\ b_1 \\ 0 \end{bmatrix} \quad (106)$$

We must check the controllability of the augmented system. Defining a matrix M_c as

$$M_c = [B^1 \quad A^1 B^1 \quad A^2 B^1] \quad (107)$$

it can easily be checked that the determinant of M_c equals zero when the matrices A^1 and B^1 taken from equation (106) and hence M_c is rank deficient. Therefore, the augmented system is uncontrollable. Since we augmented a pole at the origin and since it is uncontrollable, the system is also unstabilizable. This argument can easily be extended to a structure of larger dimension with multiple-inputs and multiple-outputs. Therefore, frequency-shaped cost functional methodology alone may not be used to design integral control for finite element models of flexible structures. It may be possible to obtain integral control when frequency-shaped LQG method is used in conjunction with low authority control [5]. The implications of this are being studied now.

8. CONCLUSIONS

Finite element models of flexible structures were considered. Frequency shaped Linear Quadratic Gaussian method was explained to design controllers for reduced order models. The most obvious benefit of frequency-weighting the cost functionals in the quadratic performance index is that it makes the evaluation model stable in closed-loop operation, while with standard LQG design it is quite often unstable. This result was clearly seen in the simple examples. The major cost of using this method is the additional states that are required in the design model. For academic type simple systems (such as the ones shown in this paper) the additional hardware needed for implementation is very simple, and the increased computational burden is minimal. However, for a realistic space structure, this addition of states would cause an end result of discarding some design states that could otherwise be kept. Therefore, a trade-off occurs: loss of model information versus reduced stability.

Another observation that has been made concerning this method is that the frequency weightings are nothing more than dynamic feedback compensators and shaping filters. By making a careful examination of the block diagrams resulting from the additional weighting states this relationship may be seen.

Application of singular value theory to analyze the stability and performance robustness of the evaluation model was discussed. For additive uncertainty, several different stability robustness tests were considered. It was shown that the sufficient condition for these tests to imply each other is that the inverse of the condition number of the loop gain transfer matrix must be unity, which cannot be satisfied for practical systems.

Several control systems were designed for a simple example and every useful design failed the stability robustness test, although the designs were stable. A suggestion to improve the applicability of the singular value robustness test is that the direction and or structure of the uncertainty must somehow be accounted for.

The use of minimum singular value plot of the loop-gain transfer matrix was emphasized since it parallels the Bode plot for single input single output systems. It was shown that integral control must be used to improve the low-frequency closed-loop performance of the finite element models when the rates of displacements are used as measurements. It was also shown that it may not be possible to obtain integral control by frequency shaped Linear Quadratic Gaussian methodology alone for structures represented by finite element models.

REFERENCES

1. Strunce R., Lin J., Hegg D., and Henderson T., "Actively Controlled Structures Theory," Final Rept, vol. 2 of 3, R-1338, Charles Stark Draper Lab, Cambridge, Mass., Dec. 1979.
2. Maitrovitch L. and Oz H., "Modal Space Control of Distributed Gyroscopic Systems" AIAA J. Guidance and Control, vol 3, no. 3, pp 218-226, 1980.
3. Sesak J.R., Likins P.W., and Coradetti T., "Flexible Spacecraft Control by Model Error Sensitivity Suppression," J. Astronautical Sciences, vol. 27, no. 2, March-April 1979.
4. Kosut R.L., Salzwedel H. and Emami-Naeini A., "Robust Control of Flexible Spacecraft," to appear in AIAA J. Guidance and Control.
5. Gupta N.K., Lyons M.G., Auburn J.N. and Marulies G., "Modeling, Control and System Identification Methods for Flexible Structures," AGARD Lecture Series AG-260 on Spacecraft Pointing and Position Control, France.
6. Skelton R.E., "Cost Decomposition of Linear Systems with Application to Model Reduction," International J. of Control, Dec. 1980.
7. Anderson B.D.O. and Moore J.B., Linear Optimal Control, Prentice Hall, 1971.

8. Strunce R., Hegg D., Lin J., and Henderson T., "ACOSS Four (Active Control of Space Structures) Theory," Final Repts, Vol 1 of 2, RADC-TR-80-78, Charles Stark Draper Lab, Cambridge, Mass., April 1980.
9. Gupta, N.K., "Frequency-Shaped Cost Functionals: Extensions of Linear-Quadratic-Gaussian Design Methods," AIAA J. Guidance and Control, Nov./Dec. 1980, pp 529-535.
10. Bullock, T.E. and Elder J.M., "Quadratic Performance Index Generation for Optimal Regulator Design," 1971 IEEE Conference, Decision Control, pp 123-124.
11. Doyle, J.C., and Stein, G., "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis" IEEE Transactions on Automatic Control, Vol AC-26, No. 1, February 1981, pp 4-16.
12. Lehtomaki, N.A., Castanon, D., Levy, B., Stein, G., Sandell, N.R., Jr., and Athaus, M., "Robustness Tests Utilizing the Structure of Modelling Error," IEEE Conference on Decision and Control, 1981, pp 1173-1190.

APPENDIX A

Spectral Factorization

Most previously proposed weighting functions have been simple ratios of pole and zero pairs, usually adaptable into the form $\omega^n + a^n$, where n must be even. These are easily factored as shown.

o If $n/2$ is even, $\omega^n = s^n$

$$s^n + a^n = 0 \quad \text{or} \quad s^n = -a^n$$

therefore $s = a \sqrt[n]{180^\circ/n + 360^\circ i/n}$
 $i = 1, 2, \dots, n$

o If $n/2$ is odd, $\omega^n = -s^n$

$$-s^n + a^n = 0 \quad \text{or} \quad s^n = a^n$$

therefore $s = a \sqrt[n]{360^\circ i/n}$
 $i = 1, 2, \dots, n$

the product of all the left-half plane roots yields the numerator or denominator of the elements in the matrices that must be realized (P_1 or P_2).

APPENDIX B

TRANSFORMING MODIFIED ALGEBRAIC RICCATI EQUATION

$$\begin{aligned} -PA - A^T P - Q + (PB + N)R^{-1}(PB + N)^T &= 0 \\ -PA - A^T P - Q + (PBR^{-1} + NR^{-1})(B^T P + N^T) &= 0 \\ -PA - A^T P - Q + PBR^{-1}B^T P + PBR^{-1}N^T + NR^{-1}B^T P \\ &\quad + NR^{-1}N^T = 0 \\ -P(A - BR^{-1}N^T) - (A^T - NR^{-1}B^T)P - (Q - NR^{-1}N^T) \\ &\quad + PBR^{-1}B^T P = 0 \end{aligned}$$

This has the form of a standard Riccati equation, as shown

$$-P\tilde{A} - \tilde{A}^T P - \tilde{Q} + P\tilde{B}R^{-1}\tilde{B}^T P = 0$$

where

$$\begin{aligned} \tilde{A} &= A - BR^{-1}N^T \\ \tilde{A}^T &= A^T - NR^{-1}B^T \quad \text{provided } R^{-1} \text{ is symmetric} \\ \tilde{Q} &= Q - NR^{-1}N^T \end{aligned}$$

APPENDIX C

Theorem: Let $A \in C^{m \times n}$ $B \in C^{n \times \ell}$. Then

$$\underline{\sigma}(AB) \leq \underline{\sigma}(A) \underline{\sigma}(B) \leq \overline{\sigma}(AB); \text{ if } n \leq \ell \quad (C.1)$$

$$\underline{\sigma}(AB) \leq \underline{\sigma}(A) \overline{\sigma}(B) \leq \underline{\sigma}(AB); \text{ if } n \leq m \quad (C.2)$$

In other words, if B has more rows than columns, (C-1) does not hold. If A has more columns than rows, (C-2) does not hold.

Proof: From the property of the minimum singular value, we have

$$\frac{\|Bx\|_2}{\|x\|_2} > \underline{\sigma}(B) \quad \forall x \in C^l \quad x \neq 0 \quad (C-3)$$

Therefore

$$\|x\|_2 < \frac{1}{\underline{\sigma}(B)} \|Bx\|_2 \quad \forall x \in C^l \quad x \neq 0 \quad (C-4)$$

From the property of the maximum singular value

$$\frac{\|ABx\|_2}{\|x\|_2} \leq \bar{\sigma}(AB) \quad \forall x \in C^l \quad x \neq 0 \quad (C-5)$$

Substituting (C-4) in (C-5) gives

$$\frac{\underline{\sigma}(B)}{\|Bx\|_2} \|ABx\|_2 \leq \bar{\sigma}(AB) \quad \forall x \in C^l \quad x \neq 0 \quad (C-6)$$

If B does not have full rank, $\underline{\sigma}(B) = 0$ and (C-1) is trivially satisfied. So we need only to consider B having rank = $n \leq l$.

Thus B is a linear transformation

$$B : C^l \text{ on to } C^n \quad (C-7)$$

and the transformation is many in one and on to C^n . This argument does not hold for $n > l$ and will show a counter example for that. Now let

$$Bx = y \quad (C-8)$$

(C-6), (C-7) and (C-8) imply that

$$\frac{\underline{\sigma}(B)}{\|y\|_2} \|Ay\|_2 \leq \bar{\sigma}(AB) \quad \forall y \in C^l \quad y \neq 0 \quad (C-9)$$

Hence

$$\underline{\sigma}(B) \max_{y \neq 0} \frac{\|Ay\|_2}{\|y\|_2} \leq \bar{\sigma}(AB) \quad (C-10)$$

which is the right inequality of (C-1).

Now let us prove that

$$\underline{\sigma}(AB) \leq \bar{\sigma}(A) \underline{\sigma}(B) \quad (C-11)$$

where

$$A \in C^{m \times n}, \quad B \in C^{n \times l}$$

and with no restriction on m , n and l .

From the property of the minimum singular value, we have

$$\underline{\sigma}(AB) \leq \frac{\|ABx\|_2}{\|x\|_2} \leq \frac{\|A\|_2 \|Bx\|_2}{\|x\|_2} \quad \forall x \in C^l \quad x \neq 0 \quad (C-12)$$

Therefore

$$\underline{\sigma}(AB) \leq \|A\|_2 \min_{x \neq 0} \frac{\|Bx\|_2}{\|x\|_2} = \bar{\sigma}(A) \underline{\sigma}(B) \quad (C-13)$$

which is (C.11).

Property (C.2) can be proved in the same way by considering $\bar{\sigma}(B^T A^T)$ and bearing in mind that $\sigma_i(A^T) = \sigma_i(A)$. This completes the proof.

To show that (C.1) or (C.2) does not hold for $n > l$ or $n > m$, respectively one only need to consider, say, $A = \begin{bmatrix} 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$.

APPENDIX D

Ignoring the noise terms in equations (3) and (4), arrange the matrices A_p , B_p and C_p as follows

$$A_p = \begin{bmatrix} A_1 & & & 0 \\ & A_2 & & \\ 0 & & \ddots & \\ & & & A_n \end{bmatrix} \quad B_p = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix} \quad C_p = [C_1 \ C_2 \ \dots \ C_n] \quad (D.1)$$

where

$$A_i = \begin{bmatrix} 0 & 1 \\ -\omega_i^2 & -2\zeta_i\omega_i \end{bmatrix} \quad B_i = \begin{bmatrix} 0 & \dots & 0 \\ b_{i1} & \dots & b_{im} \end{bmatrix} \quad C_i = \begin{bmatrix} 0 & c_{i1} \\ \vdots & \vdots \\ 0 & c_{il} \end{bmatrix} \quad (D.2)$$

The subscripts n , m , and l indicate the number of states, number of inputs and number of outputs respectively. It can be easily shown that

$$\begin{aligned} H_p(s) &= C_p(sI - A_p)^{-1} B_p \\ &= C_1(sI - A_1)^{-1} B_1 + C_2(sI - A_2)^{-1} \dots + C_n(sI - A_n)^{-1} B_n \end{aligned} \quad (D.3)$$

where

$$C_i(sI - A_i)^{-1} B_i = \frac{s}{\eta_i} \begin{bmatrix} c_{i1} & b_{i1} & & c_{i1} & b_{i2} & \dots & c_{i1} & b_{im} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ c_{il} & b_{i1} & & c_{il} & b_{i2} & \dots & c_{il} & b_{im} \end{bmatrix} \quad (D.4)$$

and where

$$\eta_i = s^2 + 2\zeta_i\omega_i s + \omega_i^2 \quad (D.5)$$

Now from equation (D.3) and (D.4) $H_p(s)$ can be written as

$$H_p(s) = \frac{s}{d(s)} N(s) \quad (D.6)$$

where

$$d(s) = \prod_{i=1}^n \eta_i \quad (D.7)$$

and $N(s)$ is an $l \times m$ matrix.

Standard LQG

State Weighting

Control Weighting

State & Control Weighting

$Q_1 = \begin{bmatrix} 0 & 0 \\ 0 & a^2 \end{bmatrix}$ $R_1 = [c^2]$	$P_1 = \begin{bmatrix} 0 & 0 \\ 0 & \frac{a}{c^2} \end{bmatrix}$ $R_2 = [c^2]$	$Q_1 = \begin{bmatrix} 0 & 0 \\ 0 & a^2 \end{bmatrix}$ $P_2 = c[1+c]$	$P_1 = \begin{bmatrix} c & 0 \\ 0 & \frac{a}{c^2} \end{bmatrix}$ $P_2 = c[1+c+d]$
No F_1 No G_1 $D_1 = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix}$	$F_2 = [-b]$ $G = [0 \ a]$ $H_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $D_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$	No F_1 No G_1 $D_1 = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix}$	$F_2 = [-b]$ $G_1 = [0 \ a]$ $H_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $D_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
No F_2 No G_2 $D_2 = [c]$	No F_2 No G_2 $D_2 = [c]$	No F_2 No G_2 $D_2 = [cd]$ $D_{21} = [c]$	No F_2 No G_2 $D_2 = [cd]$ $D_{21} = [c]$
$\hat{Q}_1 = \begin{bmatrix} 0 & 0 \\ 0 & a^2 \end{bmatrix}$ $\hat{R}_1 = [c^2]$ No N	$\hat{Q}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\hat{R}_1 = [c^2]$ No N	$\hat{Q}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & a^2 & 0 \\ 0 & 0 & c^2 d \end{bmatrix}$ $\hat{R}_1 = [c^2]$ $N = \begin{bmatrix} 0 \\ 0 \\ c^2 d \end{bmatrix}$	$\hat{Q}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & c^2 d \end{bmatrix}$ $\hat{R}_1 = [c^2]$ $N = \begin{bmatrix} 0 \\ 0 \\ 0 \\ c^2 d \end{bmatrix}$
$A' = A_p$ $B' = B_p$ $K = [k_1 \ k_2]$	$A' = \begin{bmatrix} A_p & 0 \\ G_1 & F_1 \end{bmatrix}$ $B' = \begin{bmatrix} B_p \\ 0 \end{bmatrix}$ $K = [k_1 \ k_2]$ $= [k_{11} \ k_{12} \ k_{21}]$	$A' = \begin{bmatrix} A_p & B_p \\ 0 & 0 \end{bmatrix}$ $K = [k_1 \ k_2]$ $= [k_{11} \ k_{12} \ k_{21}]$	$A' = \begin{bmatrix} A_p & 0 & B_p \\ G_1 & F_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ $B' = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $K = [k_1 \ k_2 \ k_3]$ $= [k_{11} \ k_{12} \ k_{21} \ k_{21}]$
$A_D = \begin{bmatrix} A_p & -B_p K \\ MC_p & A_p - B_p K - MC_p \end{bmatrix}$ $A_E = \begin{bmatrix} A_D & 0 \\ 0 & -B_r K_1 - B_r K_2 \end{bmatrix}$	$A_D = \begin{bmatrix} A_p & -B_p K_1 & -B_p K_2 \\ MC_p & A_p - B_p K_1 - MC_p K_2 \\ 0 & G_1 & F_1 \end{bmatrix}$ $A_E = \begin{bmatrix} A_D & 0 & 0 \\ 0 & -B_r K_1 & -B_r K_2 \end{bmatrix}$	$A_D = \begin{bmatrix} A_p & 0 & B_p \\ MC_p & A_p - MC_p & B_p \\ 0 & -K_1 & -K_2 \end{bmatrix}$ $A_E = \begin{bmatrix} A_D & 0 \\ 0 & 0 & B_r \end{bmatrix}$	$A_D = \begin{bmatrix} A_p & 0 & 0 & B_p \\ MC_p & A_p - MC_p & 0 & B_p \\ 0 & G_1 & F_1 & 0 \\ 0 & -K_1 & -K_2 & -K_3 \end{bmatrix}$ $A_E = \begin{bmatrix} A_D & 0 & 0 \\ 0 & 0 & 0 & B_r \end{bmatrix}$

Table 1. Matrices required for frequency-shaping

Case #	P_1	P_2	ζ_p	ω_{n_p}	ζ_r	ω_{n_r}
1	$\begin{bmatrix} 0 & 0 \\ 0 & 2.5 \end{bmatrix}$	[1]	0.725	0.937	-0.061	4.78
2	$\begin{bmatrix} 0 & 0 \\ 0 & \frac{3}{j\omega+1} \end{bmatrix}$	[1]	0.703	1.59	-0.058	4.73
3	$\begin{bmatrix} 0 & 0 \\ 0 & 2.5 \end{bmatrix}$	$[j\omega+0.5]$	0.558	1.27	0.011	4.94
4	$\begin{bmatrix} 0 & 0 \\ 0 & \frac{7}{j\omega+2} \end{bmatrix}$	$[j\omega+1]$	0.684	1.77	0.017	4.87

Table 2. Results for Example 1

Standard LQG Design							
$Q = \text{diag } \{0,0,0,10,10,10\}$				$R = \text{diag } \{0.01, 1.00\}$			
Kalman Filter Poles		Design Poles		Residual Poles			
ζ	ω	ζ	ω	ζ	ω	ζ	ω
0.0863	1.6706	0.2930	1.8435	-0.0204	9.1112		
0.1643	2.9428	0.2194	3.6796				
0.1151	4.1387	1.0000	0.8831; 11.252				

LQG With State & Control Weighting							
$P_1 = \text{diag } \{0,0,0, \frac{50}{j\omega+4}, \frac{30}{j\omega+3}, \frac{25}{j\omega+2}\}$				$P_2 = \text{diag } \{0.2(j\omega+2), 0.2(j\omega+2.1)\}$			
$\tilde{Q}_1 = \{0,0,0,0,0,0,1,1,1,0.16,0.1764\}$				$\tilde{R}_1 = \text{diag } \{0.04, 0.04\}$			
$N^T = \{0,0,0,0,0,0,0,0,0,0.08,0.084\}$							
Kalman Filter Poles		Design Poles		Residual Poles		Additional Poles	
ζ	ω	ζ	ω	ζ	ω	ζ	ω
0.1113	1.6907	0.2232	1.7443	0.0071	9.2495	1.000	1.3196
0.1617	2.9384	0.2944	3.3140			1.000	1.8375
0.1169	4.2093	0.2171	4.5719			1.000	3.4573
						1.000	3.6804
						1.000	4.2942

Table 3. Results for Example 2

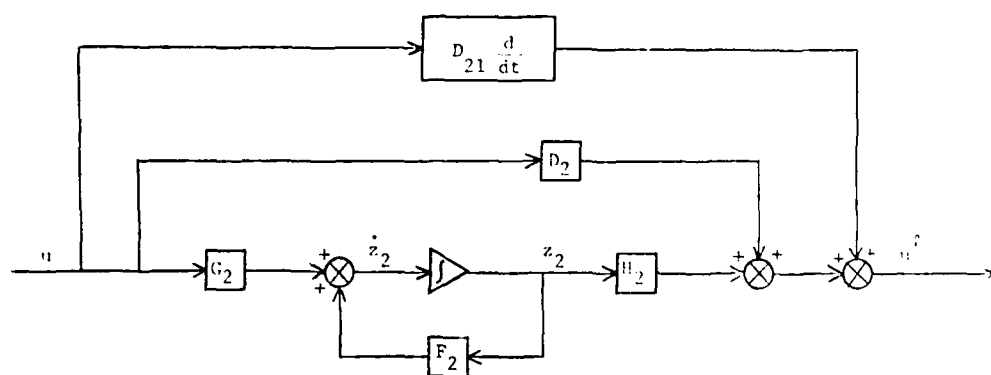
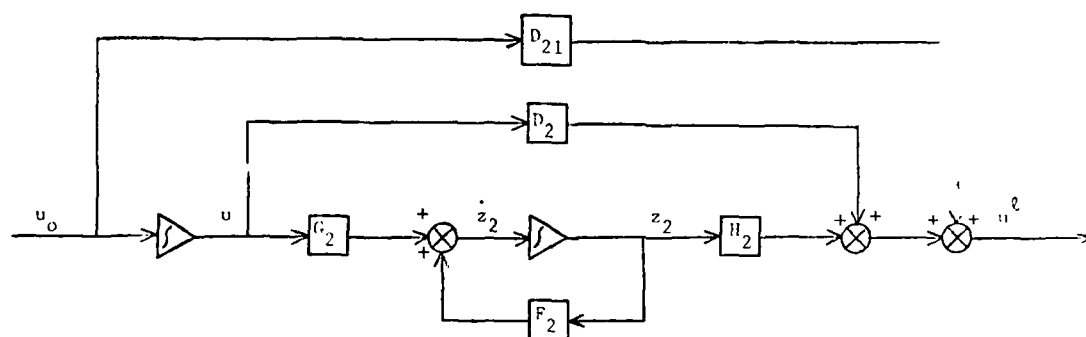
Figure 1. Block diagram when P_2 has one excess zero

Figure 2. Equivalent block diagram for Figure 1

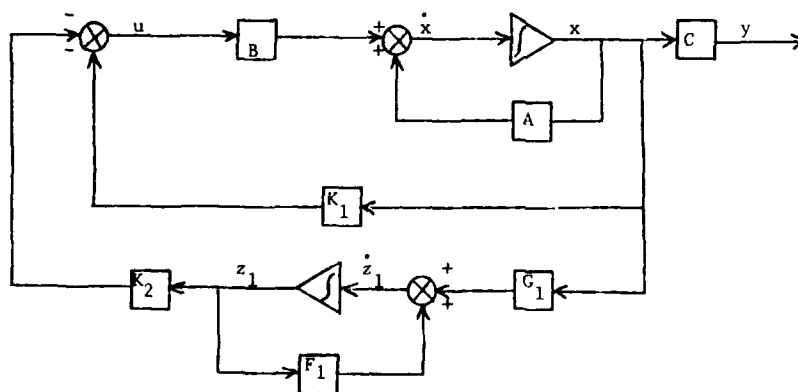


Figure 3. Block diagram for state weighting

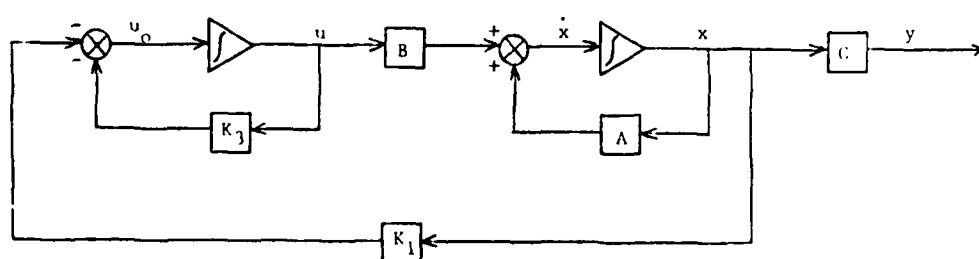


Figure 4. Block diagram for control weighting

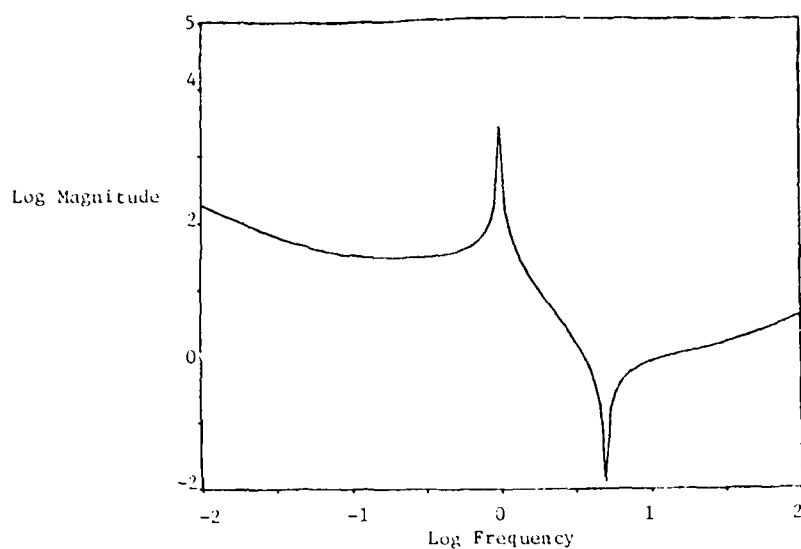


Figure 8. Stability test using rates as measurements

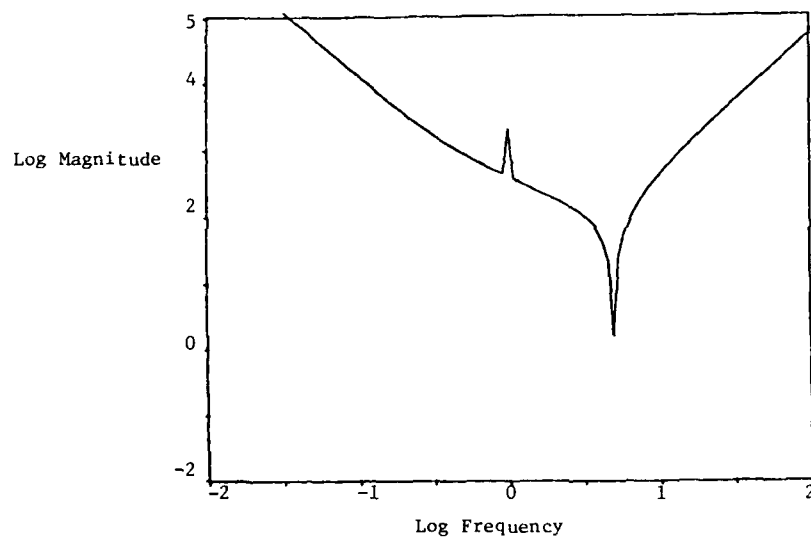


Figure 9. Stability test using rates as measurements

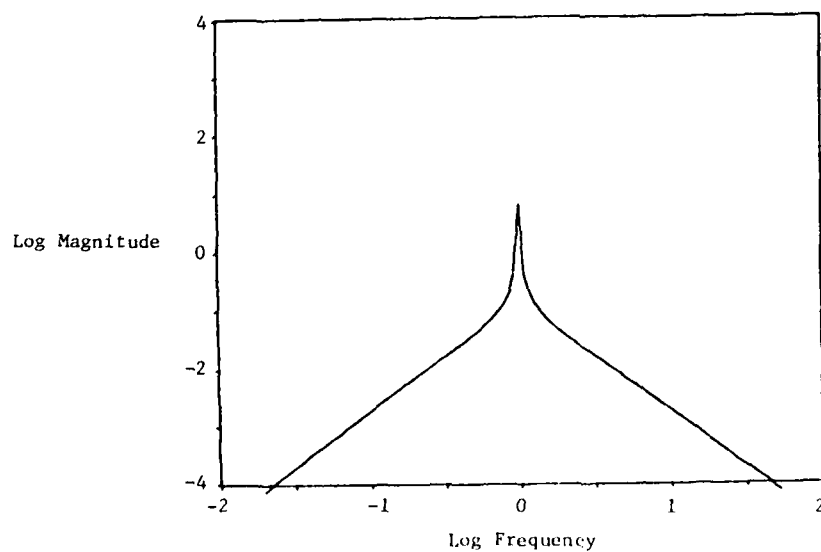


Figure 10. Nominal Loop-Gain for the design in Figure 9

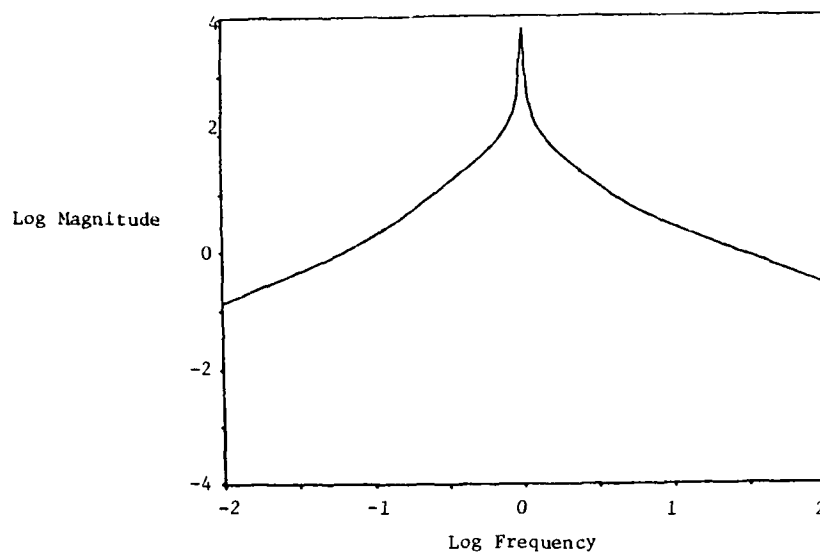


Figure 11. Nominal loop-gain for the design in Figure 8

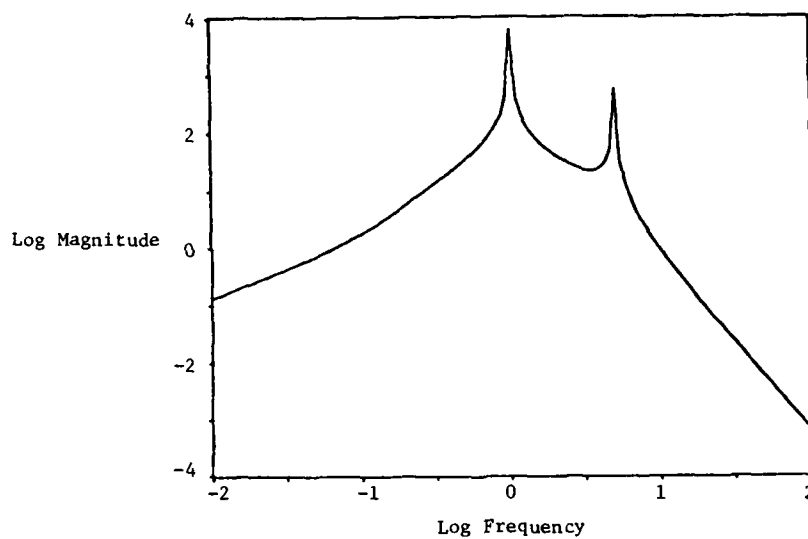


Figure 12. Perturbed loop-gain for the design in Figure 8

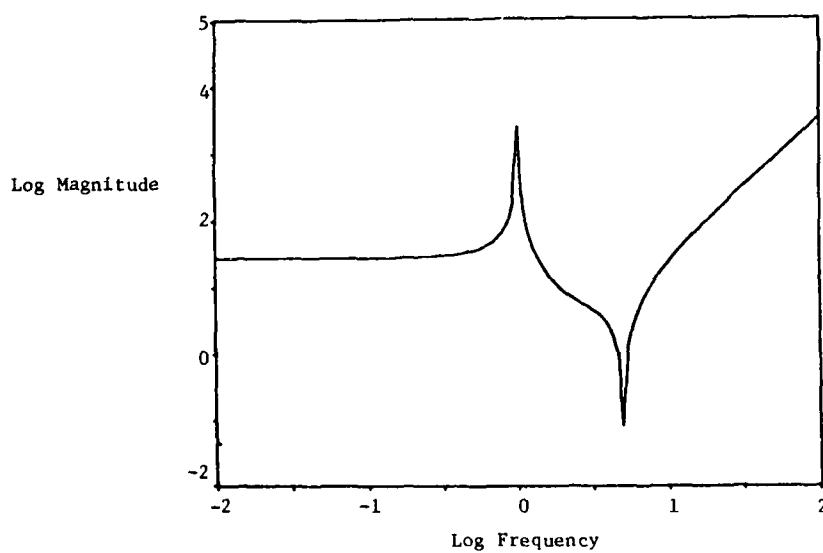


Figure 13. Stability test using displacements as measurements

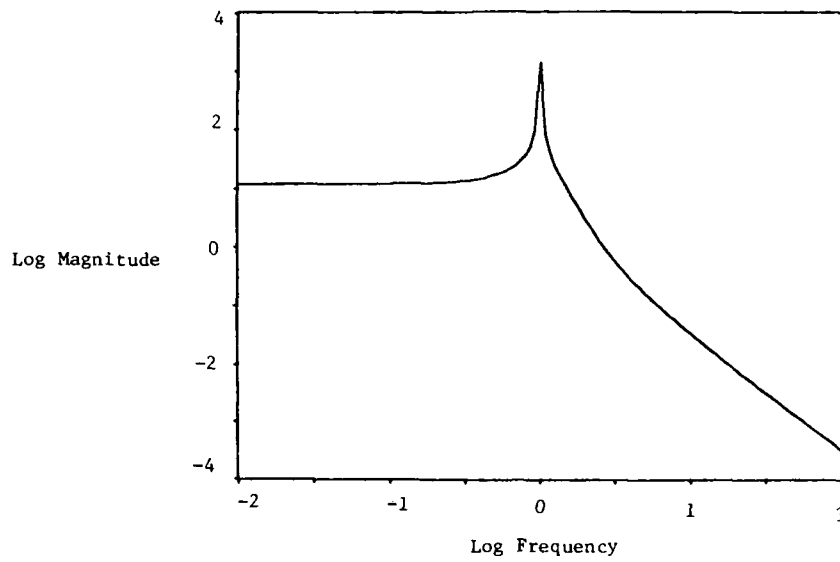


Figure 14. Nominal loop-gain for the design in Figure 13

THE DESIGN, SIMULATION AND DEVELOPMENTAL TESTING OF THE SPACE SHUTTLE DATA BUS SYSTEM

BY
ALBERT J. SHAPIRO

KEARFOTT DIVISION
THE SINGER COMPANY
150 TOTOWA ROAD
WAYNE, NEW JERSEY
07470
UNITED STATES

SUMMARY

The modern aircraft or spacecraft carrying an ever increasing quantity of sophisticated digital avionics can no longer utilize bundles of cables for interconnection purposes. Instead the time division multiplexed Data Bus System (DBS) has become the accepted technique for data transfer, significantly reducing wiring weight and complexity while increasing the integrity and reliability of signal transfer. This trend is exemplified in the B-1, F-15, F-16 and Space Shuttle Orbiter all of which use a multiplexed digital DBS.

The result of current LSI and future VHSIC component technology is to develop avionics of greater information throughput capability operating at higher frequencies yet achieving tenfold packaging densities. The aircraft or spacecraft designer faced with demands to add more mission and/or functional capability capitalizes on the above mentioned trend by adding more avionics to the vehicle thereby increasing the demand on the DBS to provide a high integrity interconnect with no degradation to the intrinsic avionic capability. The Space Shuttle Orbiter having 5 IO Processors (computers) and over 400 elements of avionics dedicated to interfacing via 24 interconnecting data buses is an example of this trend.

This paper describes a suggested sequence to be followed in the design, development, analysis and simulation of a digital DBS. It discusses the computer modeling/simulations performed and the hardware verification test results obtained on elements of the Space Shuttle Orbiter data bus. It also describes a number of specific modules developed for the Space Shuttle. It concludes with a discussion of MIL-STD-1553 Remote Terminal Bus Controllers (RT/BC) in operational use by the U.S.A.F.

1. INTRODUCTION

The designer of Guidance and Control systems for large space vehicles will no doubt use multiplexing technology for digital data communication between the elements of the system.

Multiplexing the digital bus provides many advantages: (See Figure 1)

- Minimum weight
- Fail-safe redundancy and reliability
- A high degree of immunity to spacecraft self-generated noise and external EMI
- High data rate throughput
- Standardized interface
- Ease of checkout and reconfiguration.

In addition, the nature of the DBS is such that the designer can analyze and perform computer simulation to predict system performance prior to hardware fabrication.

Recognizing the advantages, specific activity to establish a Multiplexed (MUX) Data Bus System began in early 1968 when the Aerospace Branch of the Society of Automotive Engineers (SAE) set up the Multiplexing for Aircraft (SAE-A2K) Committee made up of members from industry and the Department of Defense. The output of their efforts was a report which became the basis for USAF MIL-STD-1553 which was first issued in 1973. The Space Shuttle, B-1 and F-15 MUX developments were started before MIL-STD-1553 and were carried on during the period of its development by the SAE-A2K group. The F-16 was the first aircraft to have MIL-STD-1553 requirements applied to it.

Subsequent revisions to this military standard have been generated, the first, MIL-STD-1553A was issued in 1975 and represented a common U.S. Air Force and Navy position. As applications of MUX grew, a number of difficulties appeared in the use of this specification. A new task force was created by SAE and using inputs and comments from the three services, foreign users and industry, they issued a new document, MIL-STD-1553B in late 1978. The essential function provided by MIL-STD-1553B is to establish standards for information transfer formats and electrical interface parameters.

In the sections to follow we will describe the design sequence, technical considerations and methodology typically employed in the development, analysis and simulation of an avionic Data Bus System (DBS).

The design procedures discussed are equally applicable to applications using the MIL standard or a unique MUX system. The sequence commences with a discussion of the role of the Data Bus System designer.

2. SYSTEM DESIGN

The considerations which dictate the design choices can be divided into those dictated by the vehicle configuration (which lead to the data bus specifications) and those which constrain the specific design of the data bus itself. Figure 2 attempts to show this separation in terms of the responsibilities of the Avionics System Designer and the Data Bus Designer. The figure also outlines the key steps in the evolution of a data bus design which will be discussed in Section 3.

We will first discuss the task of the Avionics System Designer and the information which he must generate and provide to the Data Bus Designer.

Prior to the existence of MIL-STD-1553, the system designer was required to develop architectural parameters of the bus such as:

- o Signal Format
 - Waveform
 - Synch Format
- o Signal Characteristics
 - Signal Levels
 - Data Rate
 - Cable and Cable Impedance
 - Placement and Number of Stubs
- o Word Content (Number of bits, synch, parity)
 - Command Words
 - Status Words
 - Data Words
- o Hardware Characteristics
 - Impedance
 - Function
 - Response Time

The availability of MIL-STD-1553 A/B, if deemed acceptable for the intended application, significantly simplifies this initial phase inasmuch as it predetermines, by standardization, many of the essential characteristics of the DBS hardware and software.

Each DBS design is unique because each vehicle and its avionics suite is unique which still requires that the Avionics System Designer specify:

- o Number of Terminals - A detailed analysis of the bus traffic anticipated, the quantity of avionics required and master/slave relationship of the terminals is required. The degree of redundancy (as it relates to the Data Bus and the onboard avionics) must also be reviewed.
- o Physical considerations - The physical configuration of the spacecraft or aircraft must be examined to establish the dimensional characteristics of the Data Bus. The cable routing must consider:
 - a. Physical security - The cable run must be protected against accidental or hostile damage.
 - b. Temperature - Electrical and redundancy considerations may require routing into areas of severe thermal extremes limiting the ability to provide couplers/bulkhead connectors capable of operating in the environment.
 - c. EMI - Cable routing trade-offs must consider the EMI environment likely to be experienced in various areas of the vehicles. The EMI environment near powerful radar transmitters or prime electrical power generating equipment may likely impose a severe electrical problem to the equipment designer.
 - d. Vibration - The Space Shuttle data bus terminal receiver in the Solid Rocket Booster required unique mechanical characteristics to meet the vibration environment.
 - e. Electrical line length - Together with the implied consideration of bulkhead connectors, very long cable runs affect the impedance, line loss and reliability of the Data Bus. Additionally, the use of Data Bus Isolation Amplifiers may be required to extend the bus for very long runs and/or to accommodate dynamic reconfiguration of the system.

o Data Bus Characteristics.

- a. Detailed analysis is required to determine whether to use a time division command/response multiplex data bus like MIL-STD-1553 or a unique data bus protocol for the intended application. The availability of hardware and software to implement MIL-STD-1553 leaves little reason not to conform to the Standard, unless the required data/message rates or other compelling needs cannot be met by MIL-STD-1553.
- b. The Word Error Rate (WER) at specific noise levels, type of cables, noise levels and the EMI environment must be defined.
- c. The trade-offs of single or multiple clusters considering data rate prioritization, data traffic patterns and fail safe factors must be evaluated.

Allowance for expected growth in avionics is an essential consideration. Recognizing the lengthy time scale for the development of a spacecraft or aircraft weapon system and the ever changing developments of new avionics (i.e., EW, JTIDS, GPS) it would be irresponsible to not provide for planned growth in the system design of a DBS.

Another trade-off that must be considered is whether to have a single MUX terminal design to be used for all (or most) terminals. This imposes the problem of designing a MUX terminal capable of satisfying the sometimes conflicting electrical and physical requirements of many subcontractors. A common MUX terminal does enable the one time development of dedicated VLSI, common microcode and a maximized performance that is completely characterized. The use of a single MUX terminal design can and has been successfully accomplished specifically on the Space Shuttle. However, schedule constraints, GFE vs. CFE, defining clear contractual and technical responsibilities, packaging and power considerations may preclude this choice.

Based on these overall considerations, the Avionics System Designer generates the specifications for the Data Bus Designer who now must translate this specification into the Data Bus System (DBS) hardware. This is discussed next.

3. DBS DESIGN APPROACH

Having been given the System Designer's Specification, the DBS Designer must still be concerned with many facets of the DBS system design which remain unspecified. He must be cognizant of the bus structure, i.e., single level or multi-level, of the network configuration (star, ring or string bus are the most popular) since the reliability of the bus network is influenced by the topology and the control configuration (centralized, federated or decentralized) as an intertwined entity.

The typical network configurations are illustrated in Figure 3. Star configurations are frequently used where network control is located in a central node. All routing of network message traffic is from the central node to outlying nodes or from an outlying node to the central node. The star network is optimal when the bulk of communication is between the central and outlying nodes. When data traffic is high between the outlying nodes the central node is excessively burdened.

The ring configuration is arranged to form an unbroken circular configuration. Transmitted messages travel from node to node around the ring. Each node must be capable of recognizing its own address and in addition serves as an active repeater to retransmit messages addressed to other nodes. Failure of a node, adding a new node or any other break in the ring configuration will often cause the network to stop functioning. It is possible to provide for bypass of failure points but only at the cost of increased repeater complexity at each node.

The string configuration shares a single physical channel (bus) via cable linkages (bus couplers) or connectors. Messages placed on the bus are transmitted to all nodes. All nodes are required to recognize their own address in order to receive transmissions. Since the nodes are passive, network operation is inherently resistant to single point node failures.

Bus network performance is determined by bus bandwidth, number of nodes connected to each bus, bus access protocol and by system timing constraints and average and peak bus user traffic rates. This analysis is made significantly simpler for buses that are to be MIL-STD-1553B compatible. For these systems one can quickly compute the average bus loading by considering the:

- message/type
- words per message
- overhead associated with each message type
- overhead associated with node codes
- intermessage gap
- average response time
- overhead associated with non-stationary master bus controller passing.

Systems that are not MIL-STD-1553B compatible must first have established message protocol, message and word constructs and transmission rates before a similar calculation can be made.

Another aspect of data bus mechanization to be meshed with these system considerations is error management. The ability of the DBS to identify, determine the cause and achieve corrective actions of failure conditions, regardless of whether they are problems peculiar to the data bus or to an interconnecting subsystem are also part of DBS design analysis. It is the system designer's responsibility to establish the response to any error condition that may be recognized.

To achieve the proper functional partitioning of a data bus system requires the system integration knowledge, experience, and an understanding of how this can be applied to the data bus system architecture. This integration must be viewed from an overall integrated approach rather than from a conglomeration of individual sensors, controls and displays. Achieving this integration requires detail interface definitions and detailed information flow analysis. Before establishing a specific topology or control method, the detailed functional flow and message definition should be developed. It is at this time that the functional partitioning philosophy is utilized. A DBS may be partitioned by function (navigation bus, etc.) or by redundancy requirements (separation of redundant elements). Implementation of the redundancy philosophy provides improved autonomy and failure protection. However, it may create increased traffic between bus systems and therefore require more time for messages to pass from one bus to another.

It is also necessary to specify the data bus characteristics with regard to:

- cable type including specification of capacitance, twisting, shield coverage
- cable characteristic impedance (typically 70 ohms @ 1 MHz)
- cable attenuation (typically 1 to 2 dB/100 ft)
- cable termination (provision for termination of the cable at both ends with its characteristic impedance)
- cable stub requirements - either direct coupled or transformer coupled stubs are possible. The maximum stub length should be specified. Stub clusters (close spacing of stubs) should be avoided. Also the length of individual stubs should be minimized.
- impedances induced by stub linkages and by the terminal load. These impedances are critical since it is desirable to minimize reflections caused by transmission line mismatch while simultaneously maximizing the signal power to be delivered to the terminal receiver input.
- bandwidth of transformer for stub linkages to maximize waveform fidelity.

Now that the designer has defined a first cut of the electrical characterization of the DBS and the hardware elements, he can proceed to the task of analysis and simulation by developing models of each bus element. Typical models are discussed below.

- o A DBS - Figure 4 illustrates a typical string network with transformer type bus couplers located at various nodes along the bus, with bus termination resistors at both ends, and with terminal receivers located at each stub.
- o A transmission line - Figure 5A shows generalized equations for modeling a lossy transmission line with its distributed parameters. These equations describe the incremental voltage and current propagation along the line and the line characteristic impedance and propagation constant.
- o A model of a Data Bus Coupler - Figure 5B shows a bi-directional circuit which connects the data bus to a stub. The model consists of a transformer model plus series resistance, R_s , representing the internal discrete isolation resistors.
- o The Terminal Receiver (TR) (transmitter mode) - Figure 6A is an equivalent circuit of the terminal receiver when transmitting and looking back into the output transformer. The active components in the circuit model have been linearized to simplify the data bus simulation.
- o The TR Input Impedance (receiver mode) - Figure 6B shows a terminal receiver model looking into the receiver transformer with capacitive and resistive loading to account for the effects of the turned-off transmitter, filter input loading and clipping network loading.
- o The TR input filter characteristics - Figure 6C shows a mathematical representation of the frequency domain model which includes the input transformer, a 6-pole low pass Bessel filter and a 2-pole high pass filter.

The ultimate measure of data bus performance is its word error rate (WER) which is defined as:

$$WER = \frac{\text{Sync Misses} + \text{Bit Misses} + \text{Undetected Bit Errors}}{\text{Number of Words Transmitted}}$$

For the purpose of WER performance verification a Sync Miss is defined as the inability to declare a valid sync (start reception) when a valid sync is known to have occurred. An Undetected Bit Error is defined as an undetectable error of finite but very low probability which can occur despite detection of a valid sync and no Bit Miss indication.

This error rate is fundamentally dependent upon the ratio of the signal energy to the noise energy at the input to the terminal receiver. The system designer's objective is to maximize this ratio by minimizing the signal losses and the noise levels over the data bus. A reasonable design strategy is to define a design point (signal level and noise threshold) for receiver performance. The DBS designer must now adjust his design parameters to assure that a satisfactory margin of safety exists above the design point. In general, the designer will:

- limit the cable length, stub length and number of bus couplers
- increase the required transmission voltage to its largest practical level
- establish reasonable constraints on the transmitter's rise and fall time
- assure that the cable is terminated with its characteristic impedance.

Using the lumped parameter models for bus couplers, the terminal receiver, the terminal transmitter and a reasonable value for line attenuation the DBS designer must verify that the signal strength satisfies the desired margin of safety requirement at each terminal receiver for all data bus configurations. Next, the DBS designer should prepare an analysis of the expected bit error rate at the system design point. After determining that the signal strength satisfies the design requirements the concern is for waveform fidelity. Using the lumped parameter models previously developed and the distributed parameter model for the transmission line, the designer analyzes all data bus configurations and substantiates the adequacy of waveform fidelity. In this regard, experience on the Space Shuttle Program has revealed the following guidelines for improved waveform fidelity and thereby optimized WER performance:

- Avoid locating stubs in clusters (less than 1 foot apart). The problem with clusters results from reflections from other members of the cluster. These reflections tend to distort the receiver waveform.
- Avoid wide variation of stub lengths.
- Error rate is very sensitive to bus complexity (i.e., the more complex the more distortion).

At this point, the designer has verified performance and has a final electrical characterization of the DBS system and its hardware elements. Specifications for the hardware development can now be released to the MUX hardware designer who will use these DBS parameters for selecting the detector algorithm and threshold to achieve the WER.

After the hardware design details have been established the DBS designer can perform the analysis and verification of the DBS performance. This is discussed in the next section.

4. ANALYSIS AND SIMULATION

DBS performance (allowable WER) in response to nominal and worst case conditions of signal amplitude, rise time and fall time, bandwidth ambient noise are now to be analyzed and simulated. To accomplish this it is necessary to create a more complete computer model.

In addition, it is necessary to have knowledge of the terminal receiver threshold detection circuitry, filter characteristics and detection algorithms discussed earlier.

To establish confidence in the computer simulations it is desirable to build an actual breadboard model which can represent an anticipated worst case bus configuration. This model will enable, by test and measurement, the verification of computed waveform fidelity at selected terminal receiver locations and system WER performance.

A block diagram of a DBS Test Configuration is illustrated in Figure 7. Similar test configurations were built on the Space Shuttle program and proved to be vital to final refinement of the computer simulations, to establishing confidence in these simulations and also for WER screening of all production hardware. The Data Bus Controller simulates actual bus transmissions and allows for the variability of transmission parameters such as signal amplitude and signal rise and fall time. Additionally, it provides a compatible interface for the Host avionics side of the terminal receiver (MIA)* to enable the extraction of the received data and the subsequent comparison with the transmitted data. The transmitted data is routed to selected stubs in the bus configuration that is being modeled. The signal at a selected stub location is mixed with band limited Gaussian white noise in the mixer and the combined signal goes to the terminal receiver (MIA). By judicious control of signal characteristics and RMS noise levels parametric studies of WER were performed.

Figures 8A, 8B and 8C illustrate the close correspondence actually obtained between the terminal receiver waveforms generated by computer simulation and by measurement (inset on each figure) for several bus configuration and stub locations.

Once confidence has been established in the computer simulations via hardware verification, it is reasonable to perform additional computer and/or hardware studies of the effects of parametric variations of signal amplitude, system noise, terminal receiver bandwidth, terminal receiver thresholds and signal rise time on Word Error Rate (WER).

* Terminal Receiver nomenclature for Space Shuttle Applications is "Multiplex Interface Adapter" (MIA).

This was done on the Space Shuttle program and Figure 9 illustrates the anticipated performance for representative conditions. For each threshold voltage the lozenge bounds the WER expected for the given variation of signal amplitude and noise level. As shown, for measurement errors within the bounds given for signal and noise the WER rate will vary several orders of magnitude. Figure 10 compares WER results observed by computer simulation and by hardware verification testing. At any signal to noise ratio, the results correspond within one order of magnitude. This is well within the uncertainty possible because of measurement errors and the tolerance of the operating point (bandwidth and threshold) of the device under test.

Use of these computer simulations will provide advance information on the adequacy of the DBS design or of the need to make design changes in order to meet system requirements. In either event, these results are invaluable and do result in a fully defined and analyzed DBS.

5. HARDWARE ELEMENTS OF THE SPACE SHUTTLE DATA BUS

The Space Shuttle and its payload are equipped with highly integrated avionics, the whole of which is tied together by a system of 24 multiplexed data buses. Sensor outputs and controller commands are converted to digital format and transmitted as serial bit streams on the Data Bus System. The Serial Multiple Interface Adaptors (SMIA) provide the electrical transformations needed to couple the host avionics to the common bus.

The Space Shuttle was the first operational flight vehicle to utilize a digital data bus flight controls systems mechanization and a single design SMIA. There were 10 different hardware items supplied by 8 different companies which used the Singer Kearfott SMIA modules.

Aside from the shielded twisted-pair cable, the other basic element of the bus network is the Data Bus Coupler (DBC) which effects the connection between the Data Bus and the SMIA.

Another important module used in the Space Shuttle DBS is the Data Bus Isolation Amplifier (DBIA) which provides bus to bus isolation and acts as a repeater for special portions of the DBS. In particular, it services the SMIA in the Solid Rocket Booster (SRB) and allows separation of that unit without affecting the remaining DBS.

The block diagram of Figure 11 depicts these essential elements in their relative electrical relationships. The 5 IO Processors (computers) are shown interfaced to the vehicle subsystems over 24 data buses. The Shuttle subsystems are connected to the data buses by multiplex/demultiplex (MDM) units (and other terminal units) which utilized the single design SMIA for coupling to bus. The total numbers of each module of the DBS are:

Single Multiplex Interface Adapter (SMIA)	- 99
Dual Multiplex Interface Adapter (DMIA)	- 60
Solid Rocket Booster MIA (SRB-MIA)	- 8
Data Bus Coupler (DBC)	- 222
Data Bus Isolation Amplifier (DBIA)	- 2

The reader now being aware of the large number of DBS hardware elements, and the high quantity of Space Shuttle avionics being interconnected can better appreciate the critical need for a timely and detailed simulation program of the type described earlier.

Each of the DBS modules is discussed in turn below.

Serial Multiplex Interface Adapter (SMIA)

The SMIA is the interface device which provides electrical and data format compatibility between each avionics box and the common data bus. Specifically, it is contained within the particular host avionics unit LRU and provides the following functions:

- a. Signal Interchange - Receives data and command signals from the multiplex channel and transfers them to the Host subsystem. Additionally, receives data and common signals from the Host and transmits them onto the multiplex channel.
- b. Signal Format Conversion - Decodes Sync and Manchester Data to NRZ and encodes NRZ data format to Manchester format.
- c. Timing and Control - Generates the sequence signals necessary to execute, transmit, receive and transfer functions. Additionally, it accepts host-generated control signals.
- d. Signal Errors - Provides indication of parity, bit count and non-valid Manchester code errors. Design optimized for control of detected error rate of 10^{-8} and undetected error rate of 10^{-36} .

- e. Redundancy - Provides redundant control over transmitter to prevent runaway.

The SMIA is contained on a 4.4" x 2.5" x 0.4" board (see Figure 12) which is mounted within the host box. It weighs 0.3 pounds and dissipates a total of 3 watts.

Figure 13 is a block diagram of the SMIA. Its principal elements are described below.

a. Receiver

The receiver, which accepts a Manchester coded signal from the data bus, consists of a band-pass filter followed by a detector. The filter is designed to reduce signal bandwidth thereby achieving a high degree of noise rejection.

b. Transmitter

The transmitter receives data from the host LRU and converts it to Manchester code. This data is shaped to control the proper rise and fall times, then amplified and transmitted. To turn on the transmitter, all signals must be present, thus providing redundant control and fail-safe operation. These are transmitter enable and internal logic control signals. Absence of any of these signals will inhibit the output drive circuit. This prevents an output from appearing on the data bus during power sequencing and prevents transmitter run-away.

c. Decoder

The Decoder located in the EDC receives a detected signal in Manchester format from the output of the Receiver, converts it to NRZ format and transmits it to the Host.

A second output of the decoder is a 1 MHz shift clock pulse derived from the input Manchester code to control the transfer of the NRZ data to the Host.

The DMIA is a specially packaged version of the SMIA, being two MIA's on one board. This was needed to meet specific requirements of the Input/Output Processors (IOP). Each MIA of the DMIA is electrically identical to the SMIA.

The SRB-MIA's are also specially packaged versions of the SMIA. In this case a larger heat sink and a more rugged structure were used to accommodate the environment of the Solid Rocket Booster.

DATA BUS COUPLER (DBC)

The DBC unit shown in Figure 12, provides the Space Shuttle with a small, compact, data coupler specially designed, built and tested to meet the requirements of the Space Shuttle Data Bus. This unit couples the Shuttle's time-division multiplexed data and control signals between the main bus and the cable stubs connected to the MIA's in the various avionic units of the Space Shuttle.

The DBC consists of a specially designed isolation transformer, two isolation resistors of controlled resistance and inductance and a termination resistor. With these components the DBC provides impedance transformation, dc isolation, common-mode and differential noise rejection, line termination capability, and protection against shorted-stubs affecting the main bus signals. The DBC size is 1 x 1 x 0.9 in. and weighs 0.25 lbs.

DATA BUS ISOLATION AMPLIFIER (DBIA)

The DBIA shown in Figure 12 provides bi-directional communications between the Space Shuttle Orbiter Bus and both the Solid Rocket Booster (SRB) and the Ground Support Equipment (GSE). Its block diagram is shown in Figure 14. There are two isolated independent half-duplex channels each with a separate dedicated power supply. It interconnects two 300 foot party line buses and can drive a 600 foot cable to the GSE. The DBIA provides complete isolation between the Space Shuttle DBS and the SRB and GSE buses so that no perturbation is reflected back into the Space Shuttle DBS. The DBIA is contained in a 8.5" x 7" x 3" enclosure and weighs 7.5 pounds.

6. MIL-STD-1553 COMPATIBLE HARDWARE

Electrical bus parameters and the data bus waveform of the Space Shuttle when compared to MIL-STD-1553 are quite similar in many characteristics but have several significant differences. Table I details some of these differences.

TABLE I.
SIGNIFICANT DIFFERENCES SPACE SHUTTLE ORBITER VS MIL-STD-1553B

CHARACTERISTICS	SPACE SHUTTLE	MIL-STD-1553B
WORD LENGTH	28 BITS	20 BITS
GAP BETWEEN WORDS	YES	NO
OUTPUT VOLTAGE	24-30V P-P, LINE TO LINE	18-27V, P-P, LINE TO LINE
RISE AND FALL TIME	150 +50 NS	200 +100 NS
INPUT IMPEDANCE	>6 K Ω ohms	>1 K Ω ohms
INPUT COMMON MODE REJECTION	\pm 32V PK, LINE TO GND	\pm 10V PK, LINE TO GND
INPUT SIGNAL RESPONSE RANGE	0.6V TO 15V PK, LINE TO LINE	0.43V TO 7V PK, LINE TO LINE
WORD ERROR RATE (WER)	3 x 10 ⁻⁷ @ 4.0V P-P AND 300 mV RMS NOISE	10 ⁻⁷ @ 2.1V P-P AND 140 MV RMS NOISE

With the advent of MIL-STD-1553, the data bus designer is required to consider more than the simple transparent data word decoding, validation and encoding performed by the MIA as described in the previous section. A MIL-STD-1553 MUX terminal is required to function as a Remote Terminal (RT), or a Bus Controller (BC), or a Monitor or any combination thereof. Hence, not only must a -1553 MUX terminal perform the same front-end decoding, validation, and encoding, it must, in addition, decode the contents of the received command words, data words and status words to determine the required actions. These actions vary as a function of the MUX configuration, i.e., RT, BC, or Monitor. For example, if the MUX is configured as an RT and receives a command word it must;

- o Decode the command word and verify the RT's terminal address
- o Determine whether to transmit or receive the given message
- o Determine the number of words in the message.

In addition to transmitting or receiving data words, the -1553 MUX RT unit will respond with a specific status word. These protocol functions require the designer to consider the use of digital logic circuits coupled to the front-end MIA that are capable of making the required protocol decisions. The fully compatible -1553 terminal discussion that follows addresses this requirement in detail.

The present day DBS designer developing a -1553 compatible bus has available a number of architectural and hardware options. For example, Singer-Kearfott has developed a multiplexed terminal unit of universal applicability for use in avionic LRU's. The MIL-STD-1553 Bus Controller and Multiplex Remote Terminal (BC/RT) Module utilizes high density microcircuits and a flexible two-part architecture which is virtually independent of host peculiar requirements, and is therefore usable in a wide range of applications with little or no modification. It is designed to be located in the host LRU since the entire module is packaged on a single card and requires less than 10 watts in most applications.

The BC/RT can operate as a controller or as a remote terminal. The mode of operation is selectable by a Master/Slave logic signal or via the -1553 mode command for dynamic bus control transfer.

When designated as a remote terminal, the BC/RT is responsive to all -1553A and B command-response requirements. When designated as a bus controller, the BC/RT initiates and supervises all data exchanges over the dual redundant -1553 serial data bus. Data storage and retrieval at the host parallel data bus is via direct memory access.

Features include:

- a. Packaged on a single card module
- b. Performs as a Bus Controller and/or Remote Terminal capable of executing -1553B Mode and Illegal Command Word Processing
- c. Contains separable Word Processing and Message Processing/Microcontroller sections.
- d. Word Processing section:
 - o Transparent interface for -1553A or B compatibility
 - o Interfaces with dual redundant MUX buses. Performs all fast response front-end channel functions, including word validation, and command and data word detect and decode.
 - o Provides all necessary parallel data with the sense and control signals required by the Message Processing/Microcontroller section.
- e. Message Processing Microcontroller section:
 - o Transparent architecture for interfacing with various hosts including those with microprocessors
 - o Programmable
 - o 16-bit bidirectional data bus
 - o 16-bit address bus
 - o Direct Memory Access capability to interface with host microprocessor
 - o Provides memory protection and ability for indirect addressing of host main memory.

Figure 15 is a block diagram showing the two part architecture of the BC/RT and identifying the large number of functions which are packaged on the single card by the optimum utilization of LSI technology.

Communication with the host is via a 16-bit bidirectional data bus. A 16-bit address bus and direct memory access circuits perform the functions of storage and retrieval of data to-and-from host memory. Other significant circuit elements are a 16-bit terminal bus for fast response data manipulation and transfer, and the -1553 Status Word and Last Command Word registers. In addition, an on-board scratchpad register and incrementer are provided.

The BC/RT design has been packaged in various configurations for utilization in the F-111 Weapons Navigation Computer (AYK-18), B-1B Data Bus Multiplex Interface Module (MIM) and the U.S. Army's AN/ASN-137 Doppler NAV system. Figure 16 is a photograph showing each of the above mentioned terminals. Physical size varies from less than one ATR to a full ATR card dependent on complexity of the message processor, cooling characteristics and the type of componentry used (DIPS or Flatpacks).

The MIL-STD-1553B BC/RT section of the F-111 Weapons Navigation Computer successfully completed U.S. Air Force SEAFAC (System Engineering Avionics Facility) testing at Wright Patterson AFB in May of 1983.

7. FUTURE TRENDS

The next generation of avionic data buses is moving towards a mix of wire and optical fiber, utilizing data rates of 10 to 20 megabits per second in burst modes. These high data rates are essential for future aircraft, space craft or space stations, and will exploit the benefits provided by the next generation of digital VHSIC based avionic systems for timing (scheduling of events), accuracy (freshness of data), turn around times, and system-level synchronization. Computations and man-machine interface decisions will be more efficiently handled, thereby enhancing mission success.

Studies are underway by groups such as the SAE AE-9B High Speed Data Bus (HSDB) Subcommittee. These studies encompass various architecture, topology and protocol candidates. Local Area Network (LAN) architectures based on the International Standards Organization reference model, the Open System Interconnection (OSI), are providing a framework for defining the various layers of communication protocols with specific functions isolated at each level. One significant advantage to layered protocols is that differing lower-level implementation details can be hidden, while compatibility can still be achieved at the higher levels. The AE-9B HSDB Subcommittee's goal is to have a "strawman" specification by mid-'84 and a final specification by end of '86.

As high speed data buses and local area networks rapidly evolve into our future aircraft and spacecraft applications the DBS designers and data bus avionic suppliers will be challenged to create highly sophisticated analysis and simulation techniques for early verification of data bus and hardware performance. The work described in this paper will hopefully provide a useful point of departure for these future undertakings.

References

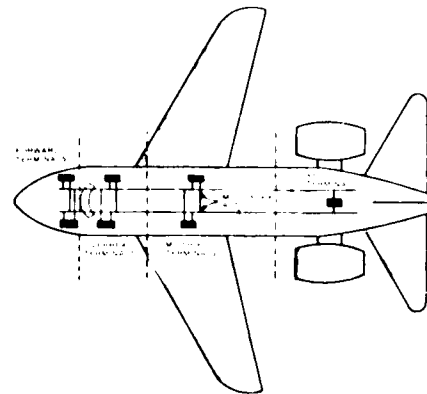
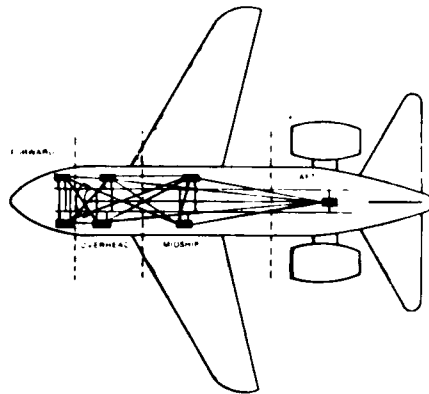
1. "Space Shuttle Multiplex Data Bus System" - William E. Mallary and Joseph E. Yeo. Proceedings Multiplex Data Bus Conference, Dayton, Ohio, November, 1976.
2. "Data Bus Simulation and Testing on the Space Shuttle Program" - Michael R. Sottile. Proceedings 2nd AFSC Multiplex Data Bus Conference, Dayton, Ohio, October, 1978.

WHY MULTIPLEXING?

AIRCRAFT WITH MANY WIRES
CONNECTING MANY SYSTEMS

VERSUS

AIRCRAFT WITH 2 WIRES
CONNECTING MANY SYSTEMS



ADVANTAGES OF MULTIPLEXING

- Multitude of interconnecting wires replaced with a very small number of twisted shielded pairs.
- Standardization of interfaces is promoted.
- System reconfiguration is immensely simplified.
- Required redundancy is built into system concept.
- System checkout can be accomplished at one point.
- Wiring checkout of the vehicle is simplified since the signals used for checkout can be the same as the operational signals.

FIGURE 1

DATA BUS DEVELOPMENT FLOW

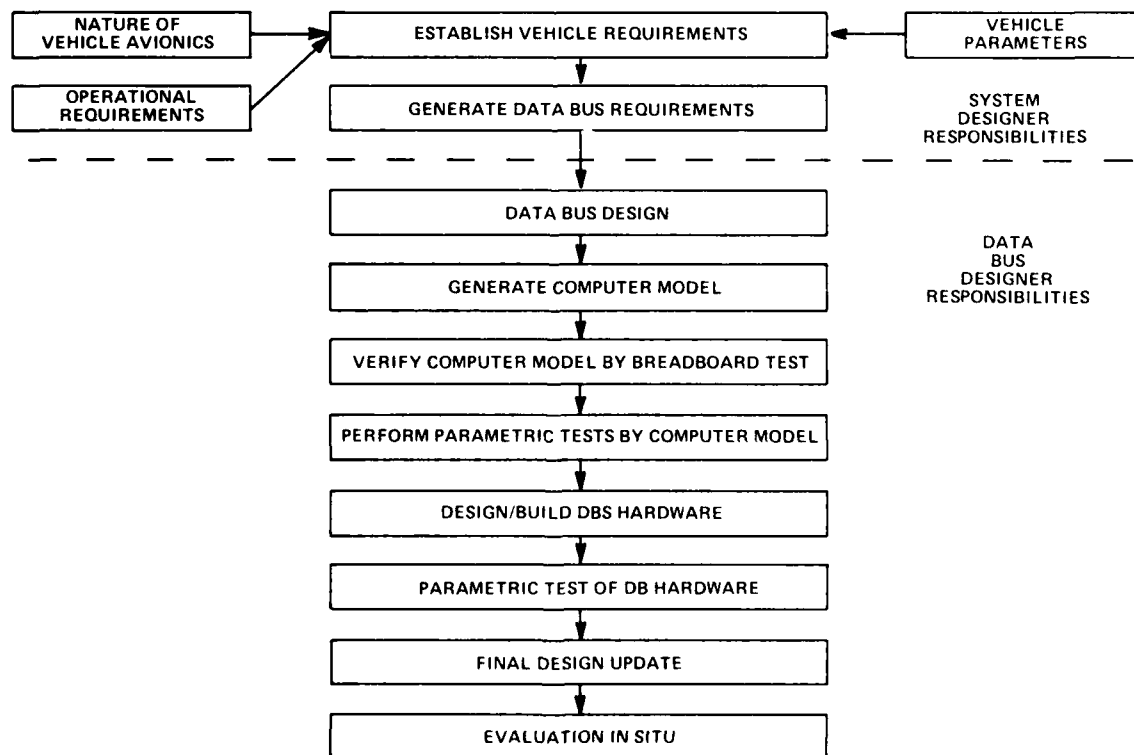


FIGURE 2

TYPICAL BUS CONFIGURATIONS

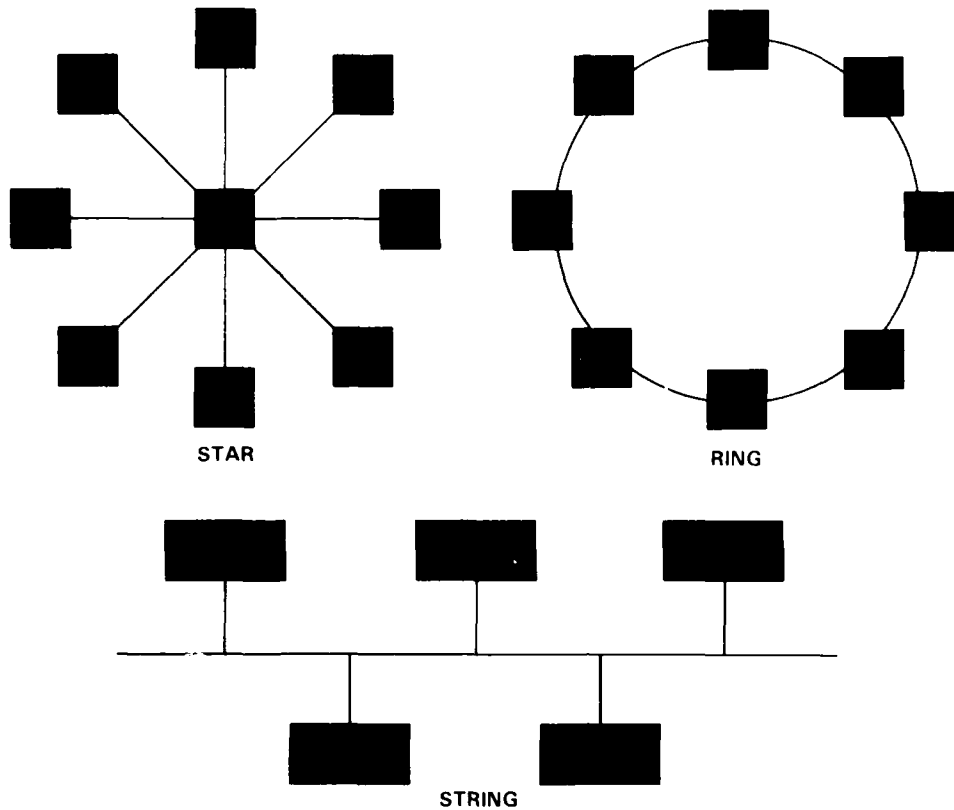


FIGURE 3

DATA BUS CONFIGURATION

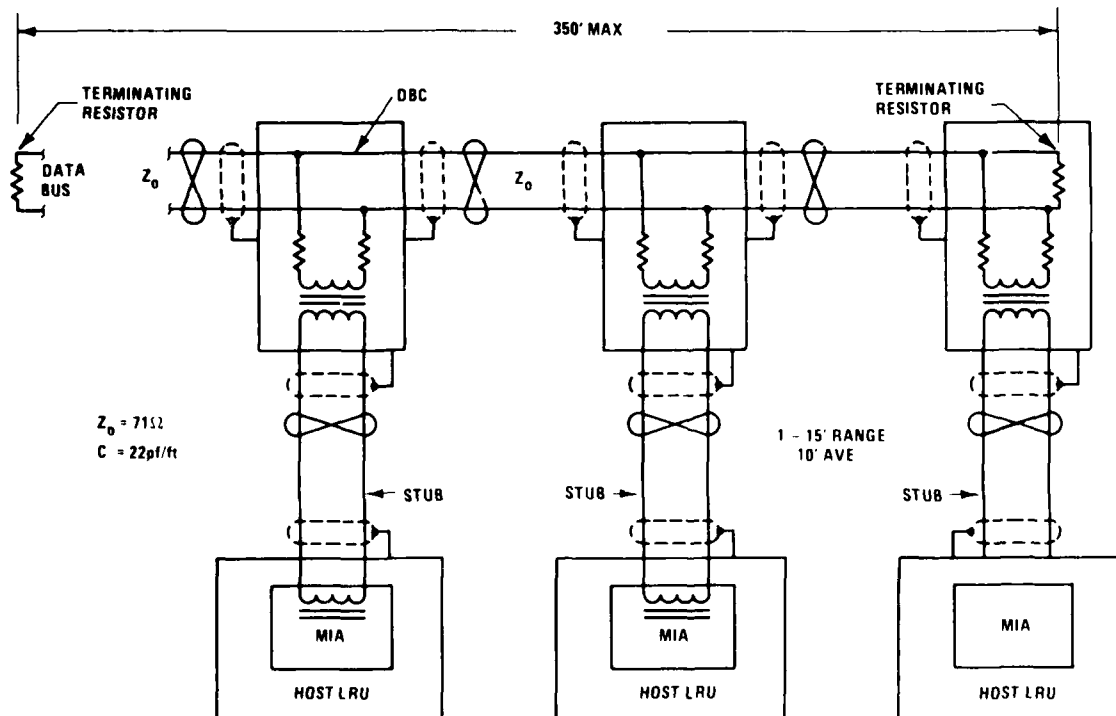
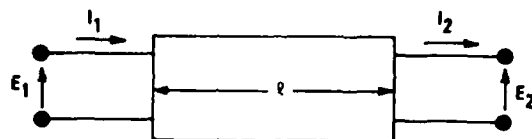


FIGURE 4

LOSSY TRANSMISSION LINE MODEL



$$E_1 = E_2 \cosh \gamma l + I_2 Z_0 \sinh \gamma l$$

$$I_1 = E_2 \frac{\sinh \gamma l}{Z_0} + I_2 \cosh \gamma l$$

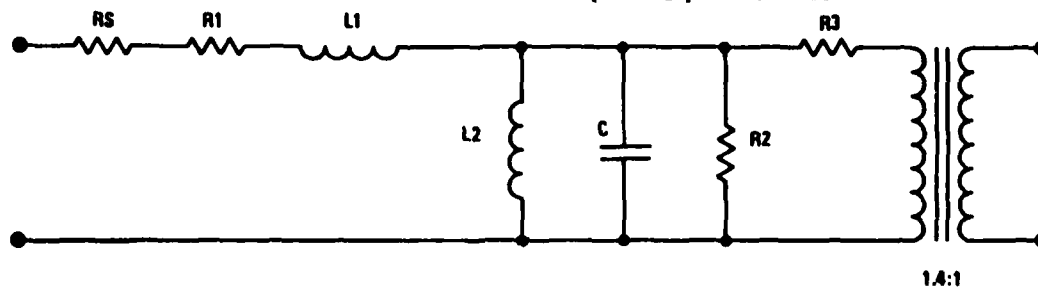
$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}$$

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)}$$

C	= distributed capacitance	pf/ft
L	= distributed inductance	nHy/ft
G	= distributed conductance	μ mho/ft
R	= distributed resistance	ohms/ft

FIGURE 5A

DATA BUS COUPLER (DBC) MODEL



RS = Internal Series Resistors ($50\Omega + 50\Omega$)

R1 = Primary Resistance

R2 = Core Loss

R3 = Secondary Resistance

L1 = Leakage Inductance

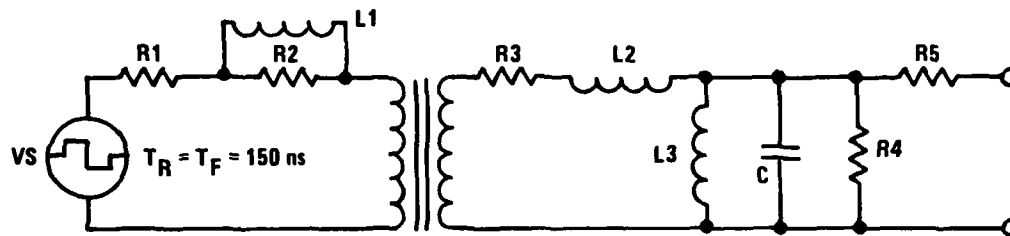
L2 = Primary Inductance

C = Equivalent Capacitance

(Includes Windings, Shield and Stray Effects)

FIGURE 5B

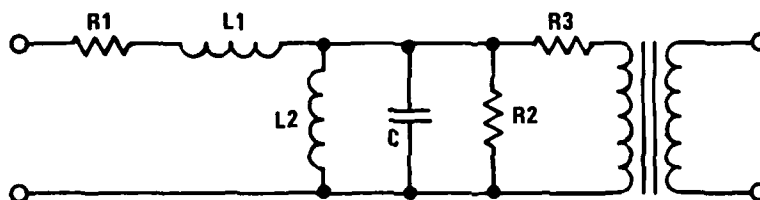
TR SOURCE IMPEDANCE (TRANSMIT MODE) MODEL



- | | |
|-------------------------------|--------------------------------------|
| VS = Source Voltage | R5 = Secondary Resistance |
| R1 = Source Resistance | L2 = Leakage Inductance |
| R2, L1 = Wave Shaping Network | L3 = Primary Inductance |
| R3 = Primary Resistance | C = Equivalent Capacitance |
| R4 = Core Loss | (Windings, Shield and Stray Effects) |

FIGURE 6A

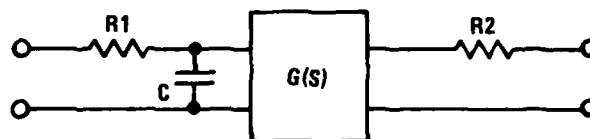
TR INPUT TRANSFORMER (RECEIVE MODE) MODEL



- | | |
|---------------------------|--------------------------------------|
| R1 = Primary Resistance | L2 = Primary Inductance |
| R2 = Core Loss | C = Equivalent Capacitance |
| R3 = Secondary Resistance | (Windings, Shield and Stray Effects) |
| L1 = Leakage Inductance | |

FIGURE 6B

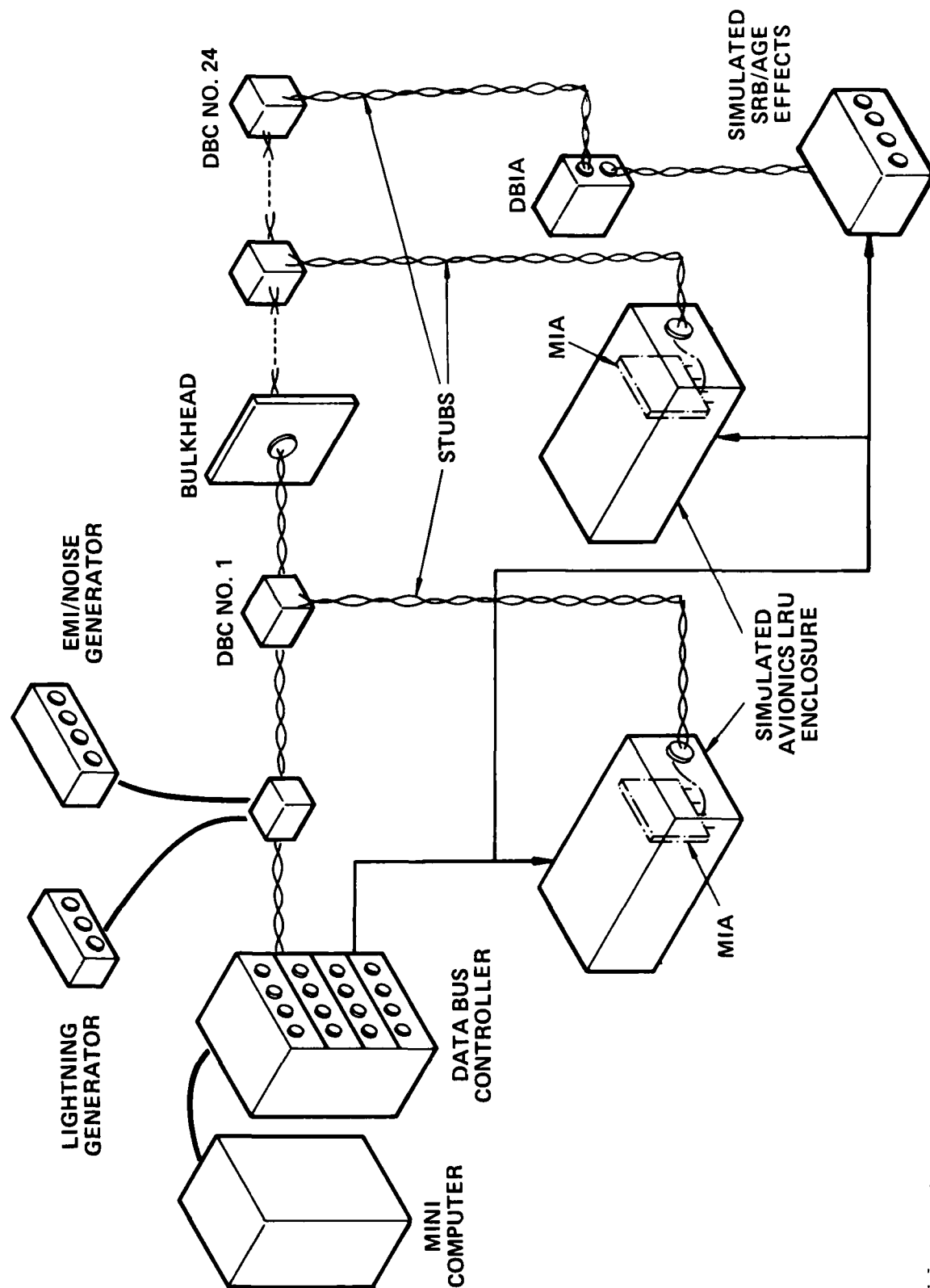
TR INPUT FILTER MODEL



- | | |
|---------------------------------|---|
| R1, C = Filter Input Impedance | $G(S) = \frac{KS^2}{S^8 + K_1 S^7 + K_2 S^6 + \dots + K_7 S + K_8}$ |
| R2 = Output Impedance | |
| G(S) = Filter Transfer Function | |

FIGURE 6C

FIGURE 7 DATA BUS HARDWARE VERIFICATION TEST SET-UP



COMPARISON OF COMPUTED AND MEASURED WAVEFORMS

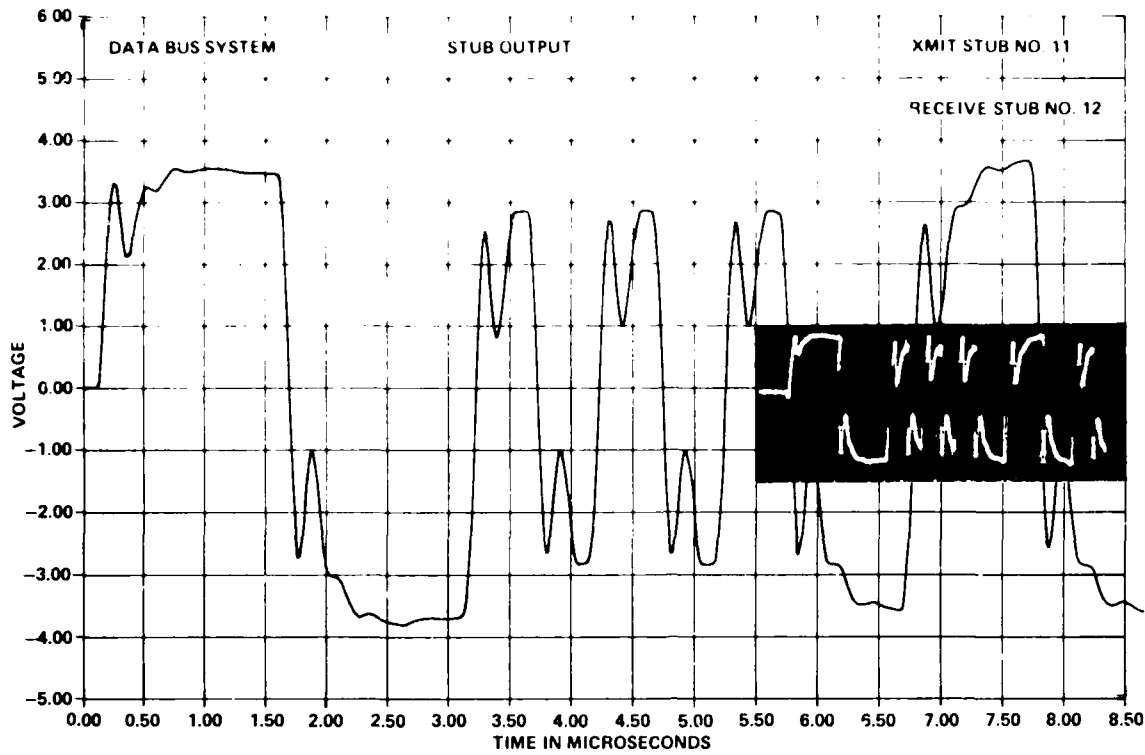


FIGURE 8A

COMPARISON OF COMPUTED AND MEASURED WAVEFORMS

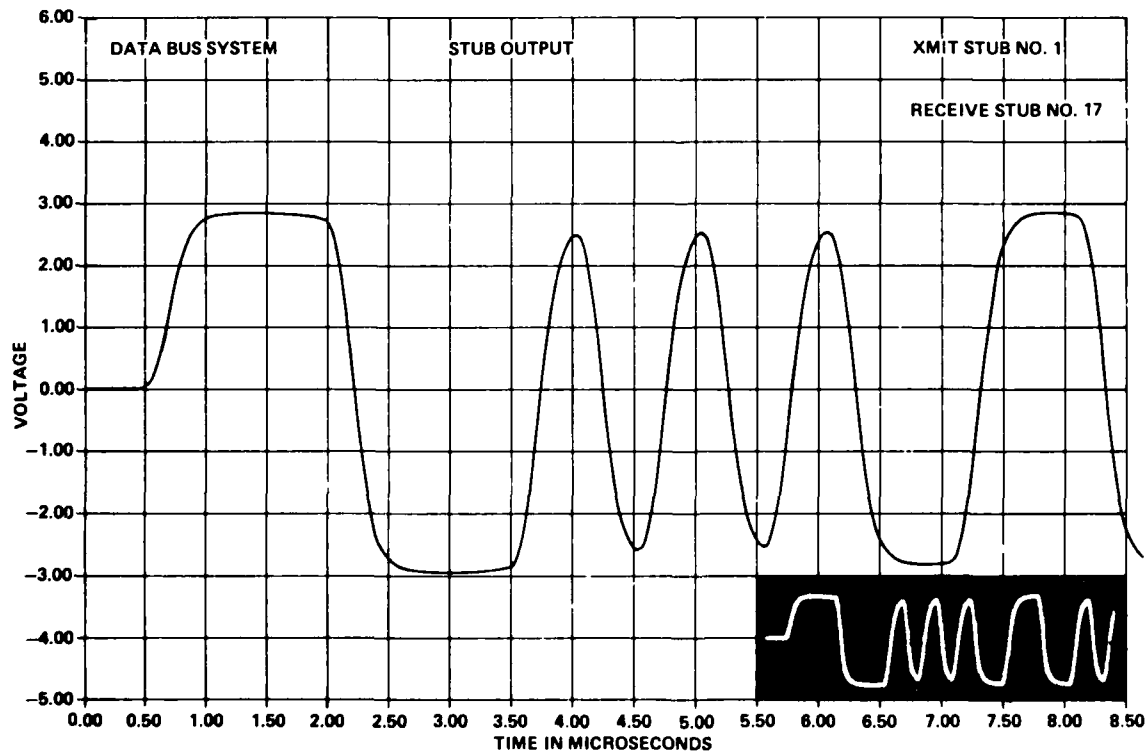
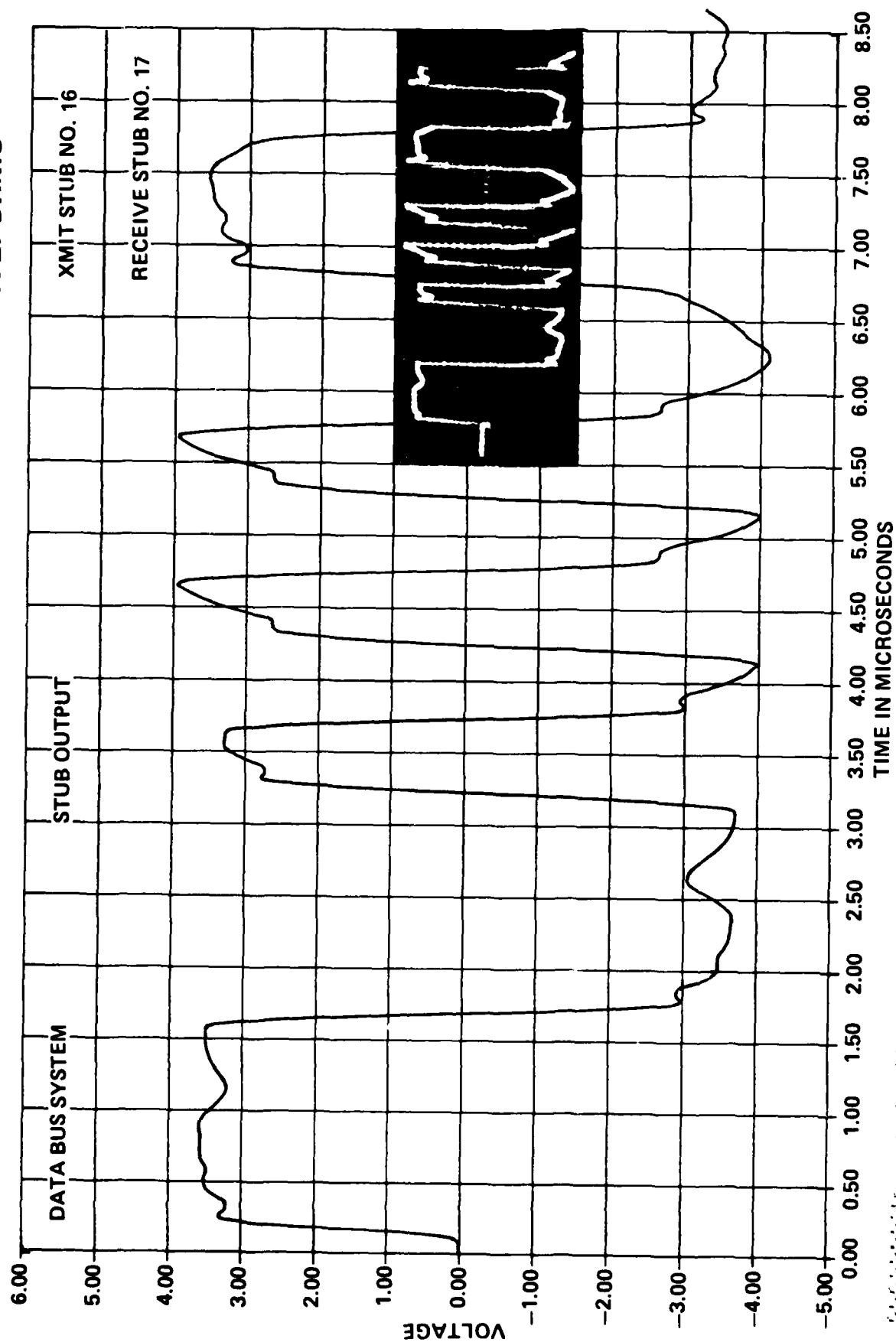


FIGURE 8B

FIGURE 8C COMPARISON OF COMPUTED AND MEASURED WAVEFORMS



WER VS SIGNAL/NOISE

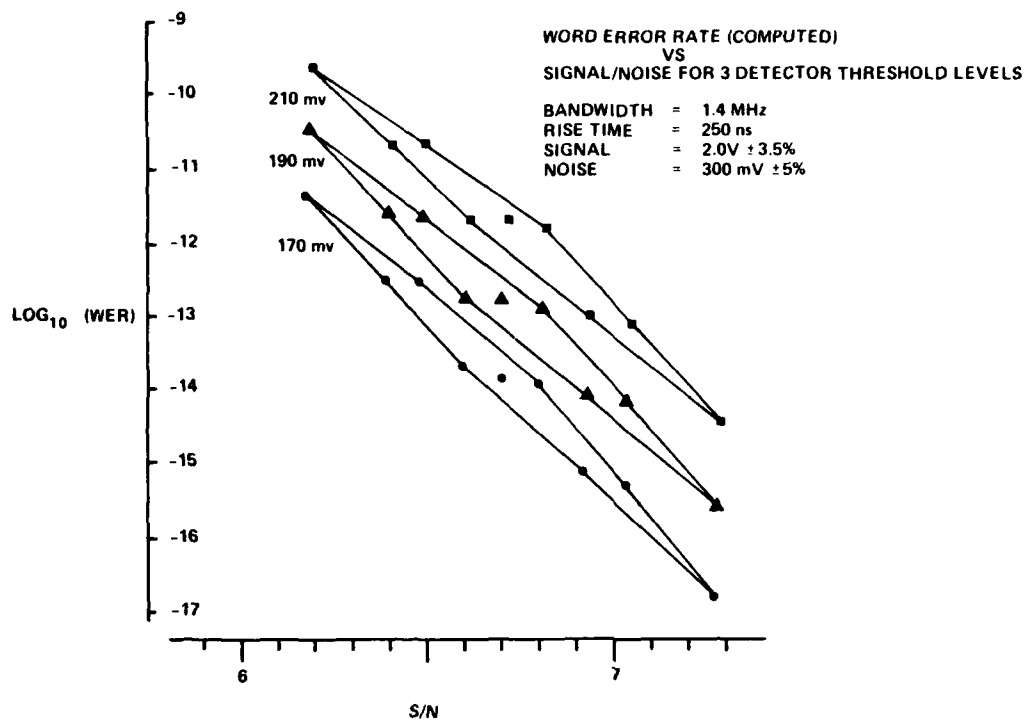


FIGURE 9

WORD ERROR RATE VS S/N

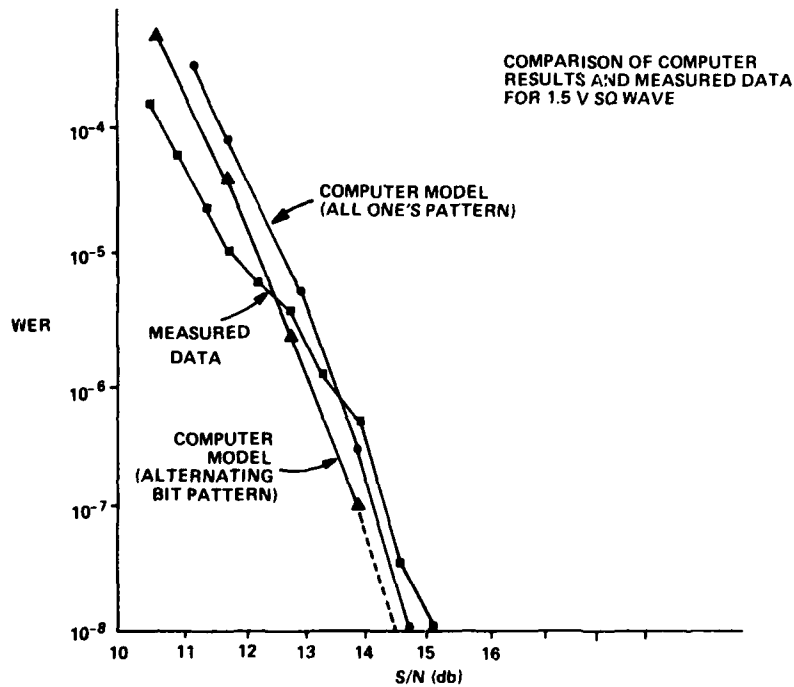


FIGURE 10

FIGURE 11 SPACE SHUTTLE MULTIPLEX BUS BLOCK DIAGRAM
(SIMPLIFIED)

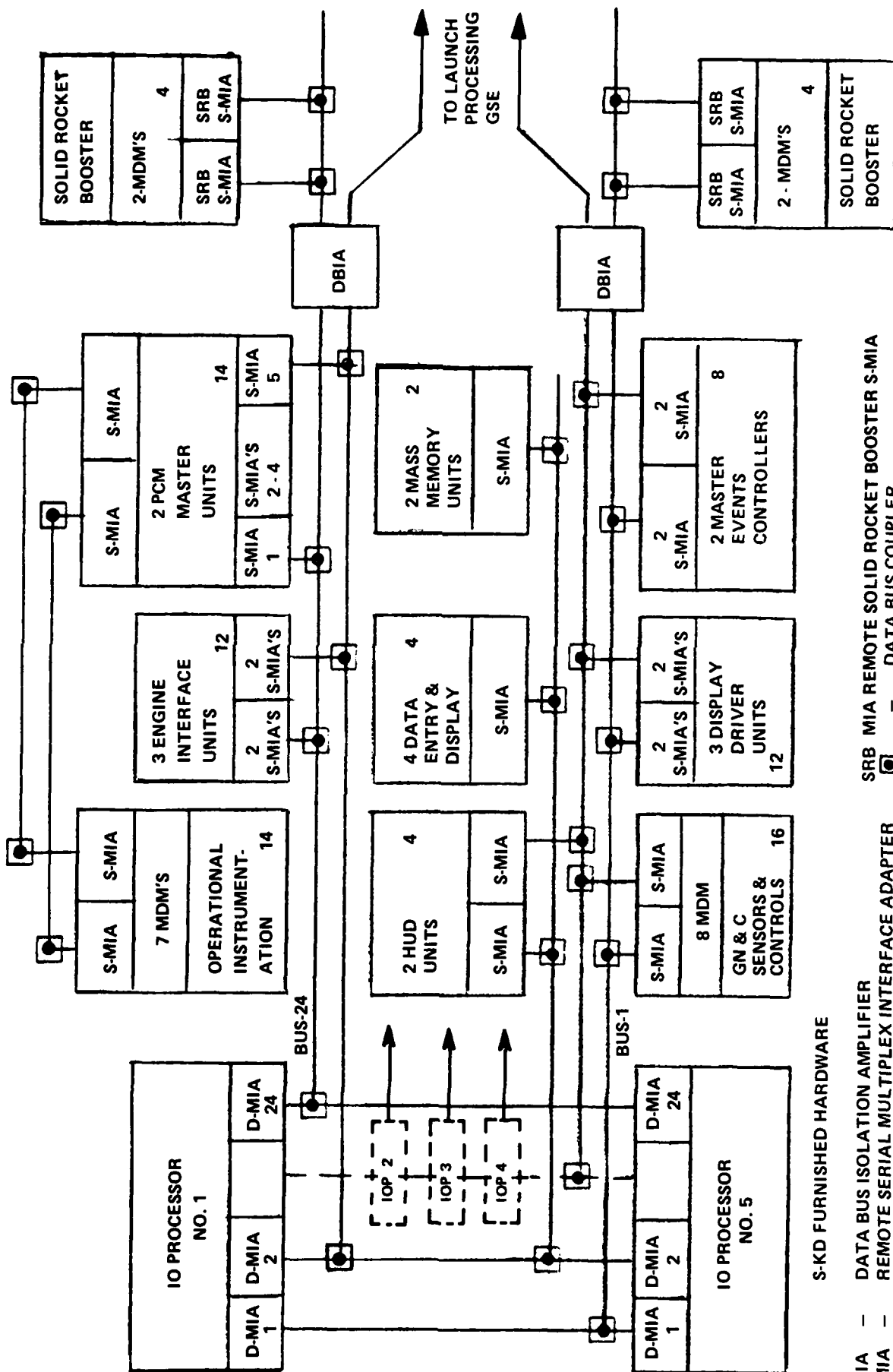
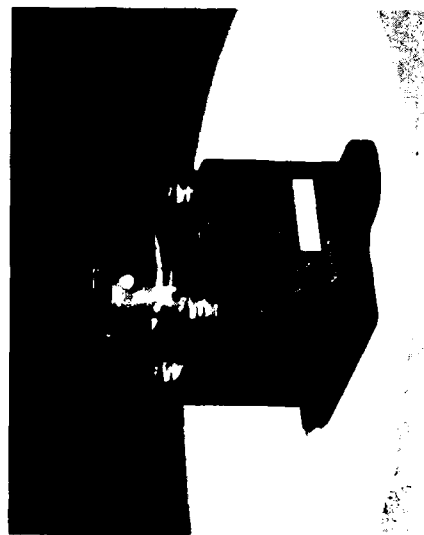
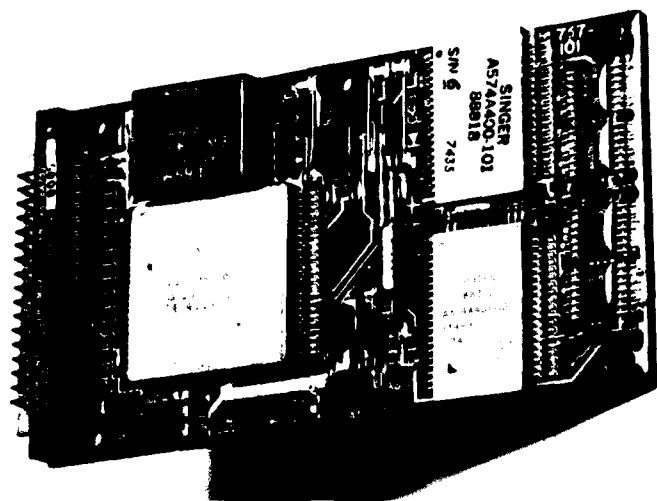


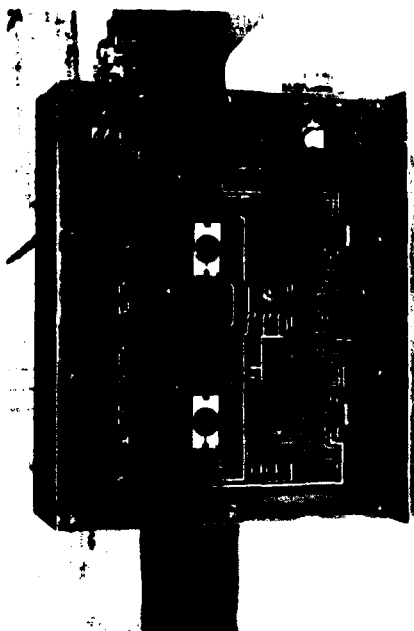
FIGURE 12 SPACE SHUTTLE EQUIPMENT



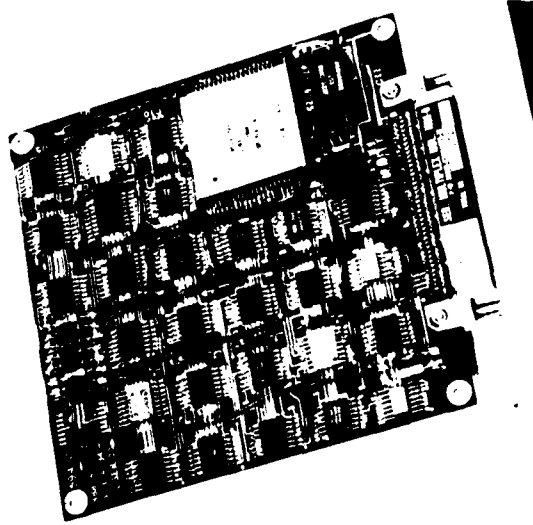
DBC



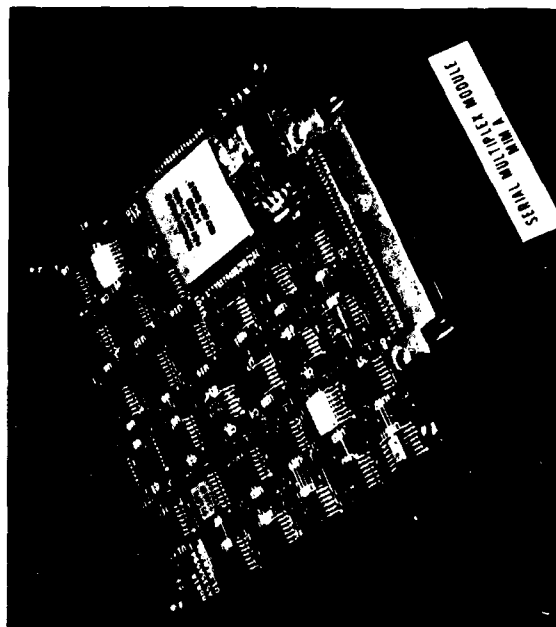
S MIA



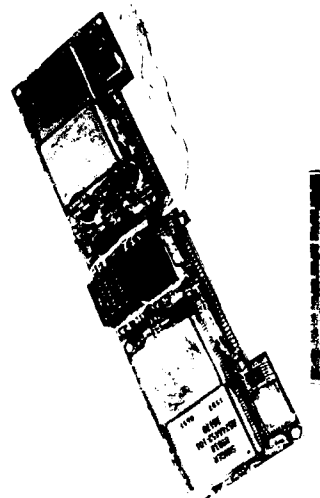
DBIA



MIM-B



MIM-A



DMIA

SMIA FUNCTIONAL BLOCK DIAGRAM

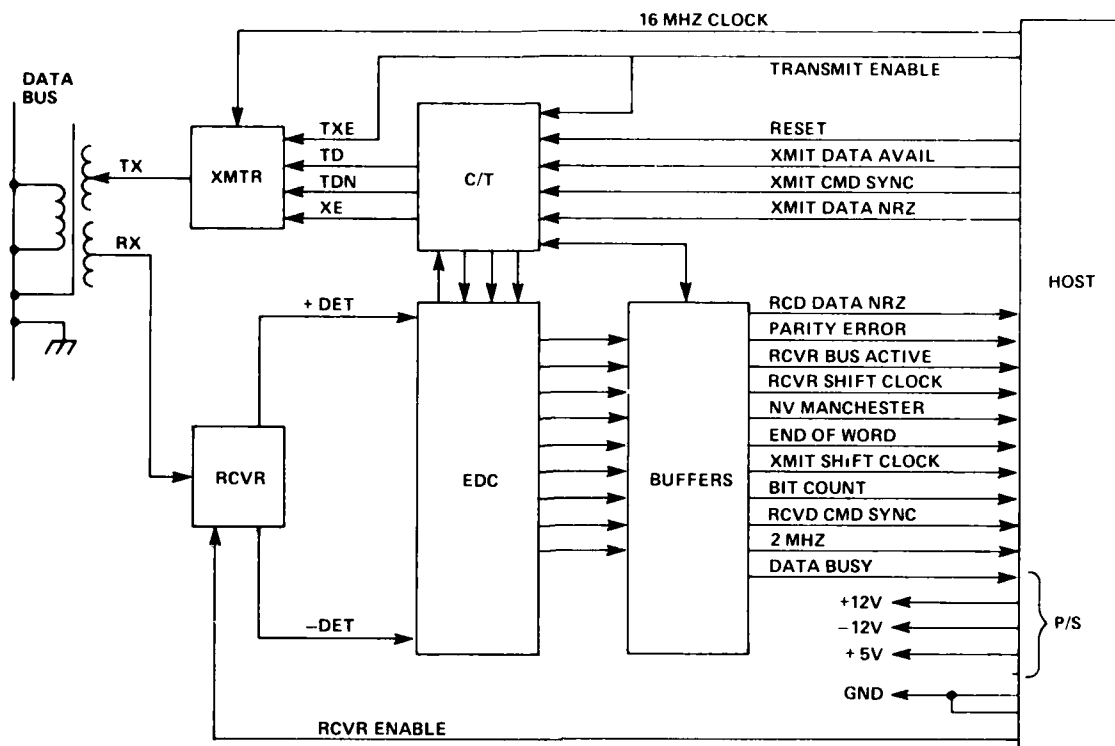


FIGURE 13

DBIA BLOCK DIAGRAM

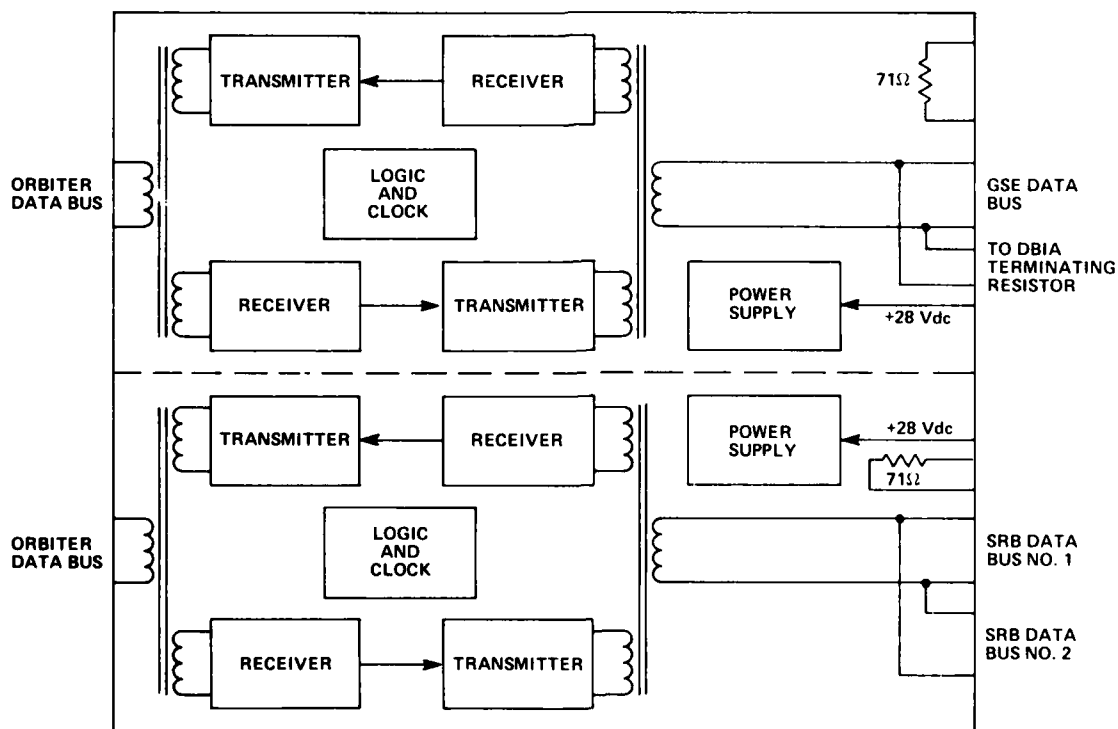


FIGURE 14

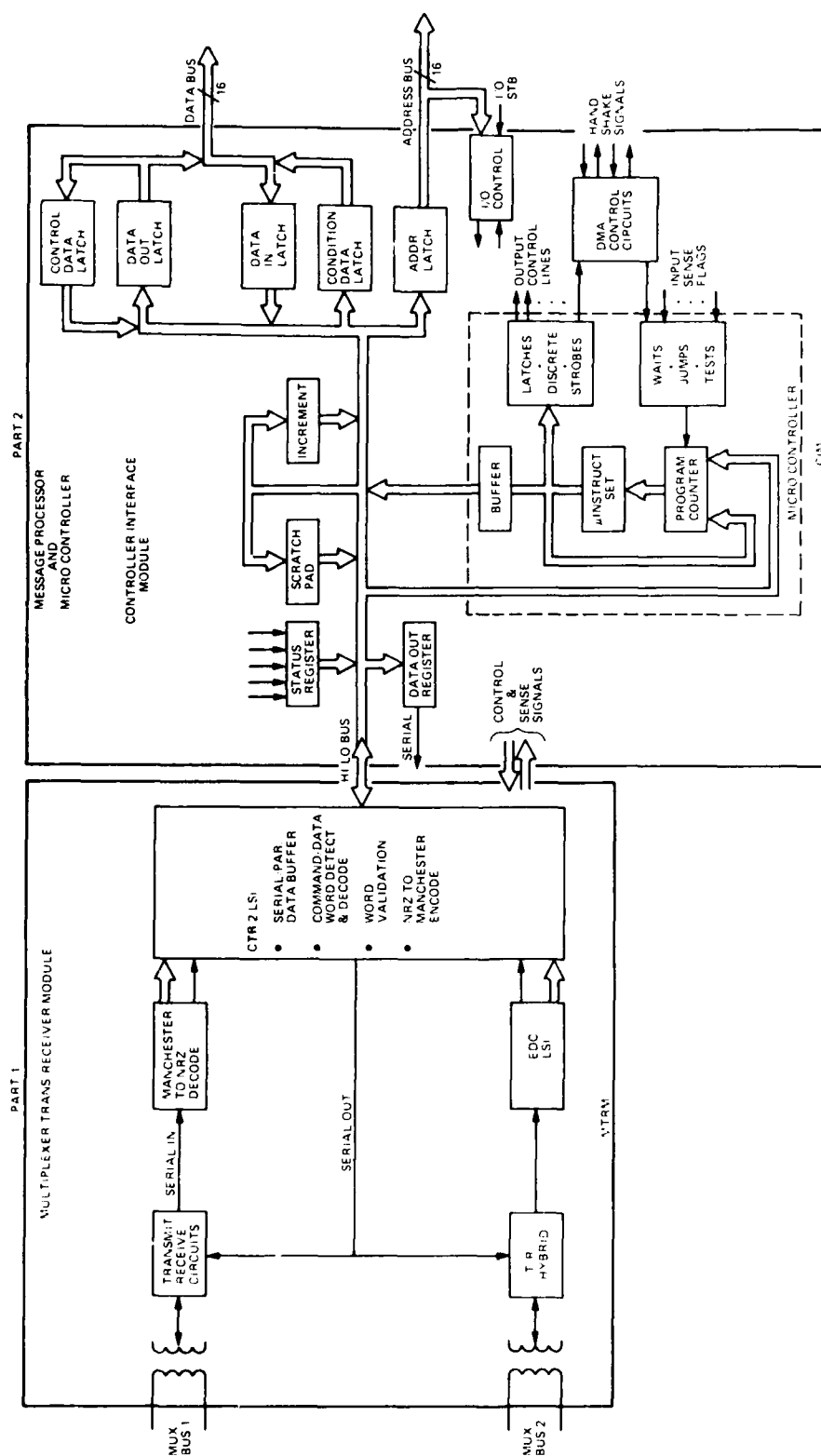
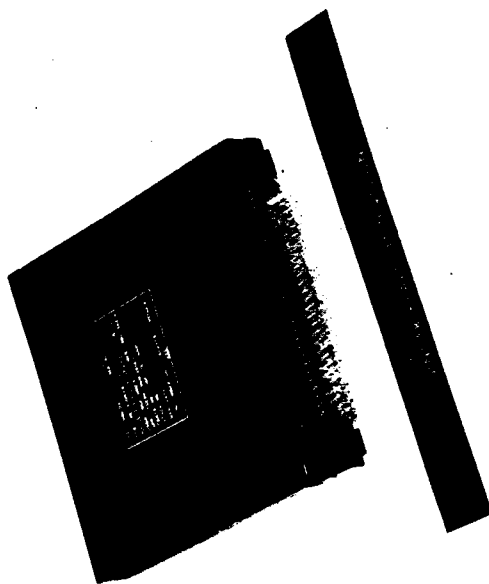
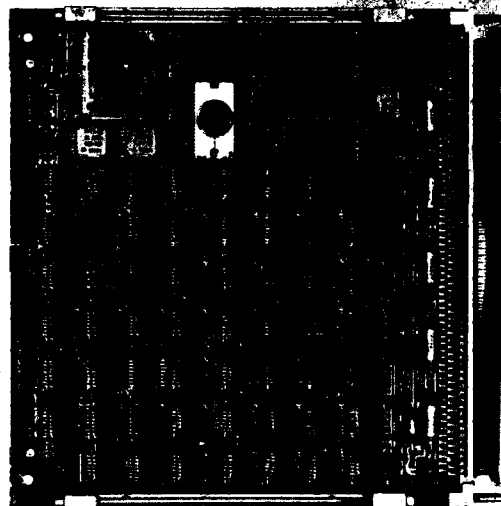


FIGURE 16 MIL-STD-1553 MUX TERMINALS



B-1B DATA BUS MIM

F-111 COMPUTER MUX



AN/ASNI-137 DOPPLER MUX

LA CERTIFICATION DES LOGICIELS EMBARQUES :

UNE APPROCHE GLOBALE ET PROGRESSIVE

G. BRACON

Electronique Serge DASSAULT
55, Quai Carnot
92214 SAINT-CLOUD
FRANCE

RESUME

La complexité et la criticité croissante des fonctions assurées par les logiciels dans des systèmes vitaux pour la sécurité engendrent un difficile problème de certification de ces logiciels. Les méthodes de travail et outils (en particulier de test) utilisés à l'ESD dans le cadre de projets tels que le Mirage 2000, ont permis jusqu'à aujourd'hui d'atteindre un niveau élevé de qualité des produits logiciels. Cependant, la croissance continue du volume des programmes et les fonctions de plus en plus vitales qu'ils exercent, exigent la poursuite de l'effort entrepris dans le domaine de l'assurance-qualité. En particulier, les techniques de vérification et de validation telles qu'elles sont mises en oeuvre actuellement sont décrites. Les études et développements en cours sont ensuite exposés : système d'aide à la définition des besoins logiciels, système d'aide aux tests dynamiques de programmes temps réel, définition des mesures de qualité qui peuvent être prélevées en s'appuyant sur un développement structuré et les produits formalisés qui en sont issus.

1. INTRODUCTION

Le fonctionnement d'un nombre toujours plus important de systèmes complexes repose sur l'utilisation des moyens informatiques et en particulier du logiciel. Ces systèmes intégrant du logiciel représentent souvent des enjeux très élevés sur le plan économique ou sur le plan de la sécurité des personnes et des nations. Tous les domaines se trouvent affectés : processus industriels, gestion des entreprises et administrations, recherche scientifique et technologique, éducation, santé, loisirs même ...! Dans le domaine des systèmes de défense, la place du logiciel augmente sans cesse. En effet, on attend l'accroissement de l'intelligence de ces systèmes, c'est à dire leur capacité à mieux exercer de façon autonome leurs fonctions de surveillance ou de destruction. Dans ce même domaine, l'apport du logiciel intervient sous le double aspect économie et sécurité. Le logiciel permet d'accroître l'efficacité des armes, donc d'en diminuer la quantité à puissance égale ; il peut aussi permettre, par la sophistication qu'il confère au système englobant, une supériorité sur les forces adverses qui est un facteur de sécurité.

Cependant, la pénétration croissante du logiciel dans des systèmes au caractère aussi vital impose qu'un degré de confiance élevé puisse être placé dans ce logiciel. Or, qu'en est-il aujourd'hui ? La plupart des logiciels existants présentent des lacunes de fiabilité : seul un fonctionnement sous contrôle humain permet leur exploitation sans risque majeur. De ce fait leur utilisation reste limitée à tous les cas où ce contrôle humain est faisable. Il en sera ainsi tant qu'il ne sera pas possible d'obtenir des éléments quantitatifs d'appréciation sur la sûreté de fonctionnement des logiciels. C'est donc en ces termes que se pose le problème de la certification des logiciels.

L'article abordera successivement les points suivants :

- Qu'est-ce que la certification des logiciels ?
- Qu'apportent les techniques de l'assurance-qualité ?
- Qu'est-ce que la "fiabilité du logiciel" ?
- L'exemple du développement des logiciels Mirage 2000
- La politique de l'ESD en matière d'assurance-qualité : une voie vers la certification.

2. LA CERTIFICATION DES LOGICIELS : UN PROBLEME OUVERT

2.1 La certification

Dans les divers domaines où des systèmes accomplissent des tâches critiques, il existe des procédures formalisées de certification. Ces procédures revêtent des formes diverses et sont identifiées par des terminologies différentes selon les contextes, c'est ainsi qu'on parle souvent aussi de qualification. Il convient donc tout d'abord de définir en général le concept même de certification, tel qu'il est entendu dans le présent article.

La certification d'un système, ou d'un sous-ensemble fonctionnellement autonome d'un système, comprend l'ensemble des procédures qui permettent de démontrer que, dans son état final, le système considéré accomplira ses fonctions sans mettre en danger la sécurité des personnes et des biens. En ce qui concerne les systèmes classiques, ces procédures ont généralement atteint un degré de maturité élevé ; en revanche, l'introduction croissante de sous-systèmes numériques n'a pas au début suscité d'aménagement des techniques de certification. Jusqu'à récemment, les logiciels intégrés aux systèmes n'affectaient pas de fonction essentielle ; la première réaction a donc été de les assimiler aux dispositifs analogiques qu'ils remplaçaient en leur appliquant les procédures habituelles. Cependant, avec l'accroissement du rôle fonctionnel du logiciel et la criticité qu'atteignent certains sous-systèmes numérisés, il est apparu qu'il fallait adapter les règles utilisées en fonction des problèmes spécifiques liés à cette technologie.

2.2 Les particularités du logiciel

En effet, le logiciel présente deux caractéristiques qui le différencient des technologies "matérielles", et qui rendent son appréhension difficile :

- La combinatoire de ses états possibles est très vite gigantesque ; ceci implique que des essais exhaustifs ne peuvent être réalisés.
- Les programmes sont immatériels ; si les textes-source représentent une description statique de ces programmes, en revanche, la dynamique d'exécution ne peut être décrite qu'au travers de documents ad hoc ; dans la pratique actuelle, aucun document ad hoc ne permet de présenter de façon simple tous les aspects dynamiques d'un logiciel temps réel de quelque complexité.

De cette seconde constatation sont nées les idées suivantes :

- Le logiciel forme un tout qui comprend les programmes et la documentation descriptive des divers points de vue : utilisation, développement, maintenance ; la présence de la documentation est en effet la condition nécessaire à la visibilité du produit.
- La production de logiciel ne peut se réduire à la programmation, mais doit inclure l'établissement de documents formalisés pour toutes les étapes du développement ; ces documents devront permettre une description exhaustive et cohérente de tous les aspects du produit final et ceci moyennant des accès faciles aux diverses informations : ce souhait légitime n'est pas encore complètement satisfait, car il engendre des problèmes techniques ardues ; une discipline spécialisée, le génie logiciel se consacre à la résolution de ces problèmes.

2.3 Erreurs et pannes

D'autre part, le logiciel présente d'autres spécificités :

- Construction purement logique, il est potentiellement "démontrable", puisque son fonctionnement n'est pas soumis aux aléas de la physique.
- De par sa nature immatérielle, il ignore l'usure et donc les "pannes".

Ces considérations, a priori favorables, doivent cependant être tempérées.

Hormis en ce qui concerne des programmes de taille très réduite et de fonctionnalité assez simple, il s'est avéré impossible à ce jour de démontrer la validité d'un code. En effet, la preuve à développer constitue en soi un travail considérable, et du fait de l'explosion combinatoire des algorithmes utilisés, s'apparente au problème du test exhaustif. Par ailleurs, la preuve de programme implique que les spécifications soient parfaitement modélisées, ce qui est rarement réalisé. Si le logiciel ne connaît pas les pannes causées par le vieillissement, en revanche du fait de sa complexité, il est souvent entaché d'erreurs résiduelles (de définition, de conception, de codage). Ces erreurs, lorsqu'elles sont déclenchées, ont des effets comparables aux pannes dans le matériel. Enfin, les programmes s'exécutent dans un environnement matériel : calculateur, moyens d'entrée-sortie ; cet environnement est sujet aux diverses sources de pannes (usure, conditions physiques de fonctionnement non nominales, erreurs de conception ou défauts de fabrication) qui peuvent avoir sur le logiciel des conséquences catastrophiques.

2.4 La certification appliquée au logiciel

De tout ce qui vient d'être dit, il apparaît que la certification du logiciel a une raison d'être, et que les procédures habituellement utilisées pour d'autres technologies doivent au minimum être adaptées.

Puisque l'exhaustivité du test de certification est exclue, il convient de combiner deux approches :

- L'application de séries de test qui sont soigneusement établies afin d'être représentatives tant du fonctionnement normal du système, que des anomalies susceptibles de se produire. [SOR 79]
- Le contrôle qu'un certain nombre d'opérations de "fabrication" du logiciel obéissent à des critères reconnus pour contribuer à la qualité du produit final (ou du moins à ceux de ces critères qui intéressent la certification, en particulier la sécurité de fonctionnement). Il faut donc orienter pour une part la certification du logiciel vers la vérification du respect de procédures au long du développement. Cette vérification ne peut pour l'instant s'appuyer que sur des critères qualitatifs en majorité. Il importe cependant de progresser rapidement dans la voie des mesures objectives, car "mesurer c'est connaître" : il n'est pas de technique mature sans métrologie établie.

3. L'ASSURANCE-QUALITE DU LOGICIEL

Le rôle de l'assurance-qualité dans la certification du logiciel ayant été identifié, il convient d'explicitier ce qu'est l'assurance-qualité, et quelle est son influence sur les problèmes liés à la certification.

3.1 Assurance-qualité et contrôle-qualité [LAW 81]

Si le concept de contrôle-qualité est bien établi dans l'industrie, celui d'assurance-qualité l'est beaucoup moins. Cependant, la littérature américaine et les performances japonaises ont mis à la mode diverses approches que l'on peut classer comme participant plus de l'assurance que du contrôle de la qualité. En effet, alors que le contrôle-qualité a pour objet la vérification après coup de la conformité d'un produit à des spécifications (fonctionnelles, technologiques, d'environnement, de normes...), l'assurance-qualité se propose de garantir et vérifier progressivement la qualité d'un produit au cours de son élaboration. [RGA 82]

D'autre part, alors que le contrôle-qualité est le fait d'une équipe spécialisée dans cette fonction, l'assurance-qualité implique le personnel productif. En effet, si les dispositions générales visant à l'obtention de la qualité relèvent d'un groupe de spécialistes, la construction de la qualité relève essentiellement de l'équipe de réalisation. La construction de la qualité est la mise en oeuvre, dans le cadre d'un projet, des dispositions générales et particulières d'assurance qualité ; elle est donc d'autant plus efficace qu'elle est totalement intégrée à la production. De ce fait, l'implication des équipes opérationnelles dans la garantie de la qualité est évidemment plus forte, et la motivation générale s'en trouve améliorée. Ce type de procédure suppose néanmoins que les opérationnels soient suffisamment formés à la qualité.

En ce qui concerne le logiciel, il faut également remarquer qu'il n'est pas le résultat, contrairement à la majorité des autres produits industriels, d'une production de série, mais au contraire qu'il s'agit en général d'une production individualisée. Par conséquent, l'idée de contrôle a posteriori s'avère peu adaptée : rentable si elle présente un caractère répétitif, l'opération de contrôle est onéreuse sur un produit unique et peut aboutir à la conclusion que le produit est mauvais, alors qu'il est sur le point d'être livré ! Il est donc clair qu'en ce qui concerne la production de logiciel, seule une procédure de contrôle "au fil de l'eau", susceptible de demander des corrections tout au long du développement, présente les garanties nécessaires à l'obtention d'un logiciel de qualité dans des conditions maîtrisables de coûts et délais.

3.2 La qualité du logiciel

Pour assurer la qualité du logiciel, il est d'abord nécessaire de la définir. Une erreur communément répandue consiste à l'assimiler à l'absence d'erreurs dans le code exécutable. Or, nous l'avons dit plus haut, le logiciel n'est pas uniquement du code, mais aussi un ensemble de documents qui le décrivent. La qualité du produit ne peut donc s'appuyer que sur celle de tous ses divers constituants. Une définition générale de la qualité peut être réutilisée avec profit : c'est "l'aptitude d'un produit ou service à satisfaire les besoins des utilisateurs". Cette définition, déontologiquement satisfaisante, demande toutefois à être précisée ; s'agissant d'un produit par nature complexe et immatériel, et de surcroît coûteux à réaliser dans l'état actuel de l'art, il convient de se limiter aux "besoins spécifiés des utilisateurs". En effet, le besoin en la matière est chose floue, donc modulable en fonction de critères contradictoires tels : les moyens financiers, les délais nécessaires de mise à disposition, la pérennité envisagée, l'appréciation de la criticité... S'il n'est pas toujours possible d'attendre d'un utilisateur qu'il formule spontanément tous ces critères, il appartient en revanche au réalisateur de les révéler autant que faire se peut, afin que la définition du logiciel à réaliser décrive ce sur quoi les parties s'accordent pour contracter.

Quelques travaux intéressants ont été effectués sur les mesures de la qualité du logiciel. En particulier, J.A. Mc Call a défini des facteurs, critères et métriques qui permettent l'évaluation de la qualité. Un tel système de mesure peut permettre la spécification d'exigences de niveau de qualité. Si les travaux actuels ne sont pas encore directement utilisables, ils constituent cependant une solide base de réflexion sur laquelle s'appuie l'ESD pour ses propres études.

3.3 Les missions de l'assurance-qualité

L'assurance-qualité du logiciel au sein d'un organisme comporte trois volets successifs :

- Définition des règles générales d'assurance-qualité
- Mise en place d'un plan-qualité pour un projet donné
- Procédures de "Suivi de la qualité" tout au long du projet.

a) Règles générales d'assurance-qualité (manuel-qualité)

Ces règles décrivent le fonctionnement global des procédures d'assurance-qualité. Elles définissent en particulier le ou les cycles de vie possibles du logiciel (c'est à dire le découpage en étapes), et les responsabilités impliquées dans les diverses procédures. Elles peuvent aussi établir des standards, que ce soit en matière de documentation, de méthodes, techniques ou outils de développement à utiliser.

b) Plan-qualité

En fonction des besoins et contraintes exprimés par l'utilisateur et du manuel-qualité cité ci-dessus, le responsable de projet et le groupe qualité de l'organisme chargé du développement rédigent un plan précis des différentes actions destinées à garantir l'obtention de la qualité

requis pour le produit considéré. Ceci peut comporter le choix parmi les alternatives offertes dans le manuel-qualité, mais également l'ajout d'actions plus détaillées et de normes à respecter. En outre, une planification et des responsabilités précises doivent être établies.

c) "Suivi de la qualité"

Ce suivi est connu dans la littérature anglo-saxonne sous le sigle de "V & V" (vérification et validation). Il regroupe l'ensemble des opérations d'évaluation qui ont été prévues dans le plan-qualité.

- Les vérifications

Les vérifications sont les opérations qui consistent à s'assurer qu'un produit intermédiaire ou final du développement est conforme aux exigences qui lui ont été assignées dans l'étape antérieure, ainsi qu'aux règles décrites dans le plan-qualité.

Elles reposent essentiellement sur la tenue de revues de projet, qui, selon leur degré d'importance, peuvent être conduites en présence de représentants du client et de membres du groupe-qualité, ou restreintes à certains membres de l'équipe de projet. Bien qu'elles puissent (et doivent de plus en plus à l'avenir) s'appuyer sur des éléments quantitatifs, elles sont par nature qualitatives. Les vérifications s'effectuent sur les descriptions successives du logiciel (définition, conception globale, conception détaillée, code).

- Les validations

Au contraire des vérifications, les validations s'attachent à l'aspect fonctionnel des produits, et non à leur aspect formel. Elles portent donc essentiellement sur les programmes et consistent à vérifier leur conformité vis-à-vis du niveau de spécification correspondant (conception détaillée pour le module, spécification fonctionnelle pour la fonction...). Ces opérations consistent essentiellement aujourd'hui à effectuer des tests de programme.

3.4 Qualité et certification du logiciel [HOW 82]

L'évaluation de la qualité a un domaine d'application beaucoup plus large que la certification. La qualité du logiciel s'apprécie sur ses caractéristiques d'exploitation, de maintenance et de transport. Il est évident que maintenance et transport sortent du cadre des préoccupations de la certification, qui ne s'intéresse par ailleurs qu'à quelques aspects des caractéristiques d'exploitation.

Si l'on reprend la terminologie de Mc Call, la qualité d'un logiciel peut se caractériser par des "facteurs". Ces facteurs sont les caractéristiques appréhendables de l'extérieur, par un utilisateur du logiciel.

Au plan de l'exploitation du logiciel, cinq facteurs sont identifiés : validité, robustesse, efficacité, maniabilité et sécurité. L'efficacité et la maniabilité qui sont des facteurs respectivement économique et ergonomique, sortent du champ d'intérêt de la certification. En revanche, validité, robustesse et sécurité sont des facteurs qu'il convient d'utiliser dans le cadre d'une procédure de certification puisqu'ils conditionnent ce qu'on peut appeler ici la "fiabilité" du produit.

4. APPROCHES DE LA FIABILITE EN LOGICIEL

Ainsi que nous l'avons dit plus haut, le logiciel ne connaît pas l'usure. Les concepts fiabilistes traditionnels ne lui sont donc pas applicables. Pourtant, il est souhaitable de pouvoir évaluer a priori la probabilité de bon fonctionnement d'un logiciel. Ceci nécessite une approche spécifique et l'emploi d'une terminologie bien définie.

4.1 Le concept de fiabilité de logiciel [TRO 79]

La fiabilité peut-être définie en ces termes : "probabilité qu'un dispositif accomplisse une fonction requise dans des conditions d'utilisation et pour une période de temps déterminées".

Du fait de la non-dégradation dans le temps, cette caractéristique ne peut être prévue pour le logiciel en fonction de statistiques sur les temps de survie de ses constituants. Il s'ensuit qu'il est nécessaire de s'appuyer sur des grandeurs mesurables et fortement corrélées à la probabilité finale de bon fonctionnement. L'état de la recherche a identifié de telles grandeurs : qualité des documents de spécification et de conception, complexité du code, taux d'exhaustivité des tests, nombre de défauts résiduels.

Compte tenu de ce qui vient d'être dit, la fiabilité du logiciel peut être définie comme suit : "Aptitude à accomplir sans défaillance chaque fonction spécifiée, avec la précision requise, dans des conditions d'utilisation données et pour une période de temps déterminée".

La notion de défaillance étant susceptible de varier d'un utilisateur à l'autre, il importe que chacun détermine le niveau de gravité qu'il donne à un type de défaillance ; ceci peut se faire par la définition de classes de criticité [DO 178]

Dans le cas d'un logiciel multi-fonctions, il se peut que les niveaux de criticité affectés aux différentes fonctions soient distincts, c'est pourquoi il importe de considérer isolément chaque fonction. La précision se réfère à la mise en oeuvre d'algorithmes numériques, dont l'instabilité ou la dérive dans le temps constituent des causes de non-fonctionnement.

Enfin les conditions d'utilisation doivent également être précisées, afin que le logiciel comporte les précautions nécessaires à sa protection face aux défaillances et agressions de l'environnement (matériel, logiciel, humain).

4.2 Fiabilité, Validité, Robustesse

Parmi les divers facteurs de qualité susceptibles d'être attribués à un logiciel, deux sont particulièrement liés à la notion de fiabilité : la validité et la robustesse. La validité se définit comme étant l'aptitude d'un logiciel à accomplir, avec satisfaction, chaque fonction spécifiée pour laquelle il a été prévu et ce dans des conditions d'utilisation données. La robustesse se définit comme étant l'aptitude d'un logiciel à poursuivre, avec plus ou moins de satisfaction, chaque fonction spécifiée pour laquelle il a été prévu, et ce dans des conditions d'utilisation non nominales provoquant des perturbations auxquelles il est censé résister. Etant donnée la définition que nous avons énoncé plus haut de la fiabilité du logiciel, un logiciel fiable doit être à la fois valide et robuste. La validité donne au logiciel l'aptitude à accomplir les fonctions spécifiées dans des conditions nominales, la robustesse fournissant l'aptitude nécessaire pour résister aux agressions prévisibles de son environnement. Cette conclusion est en accord avec les deux approches qualitatives que l'on trouve dans la littérature et qui permettent de construire la fiabilité d'un logiciel tout au long du cycle de développement. La première approche concerne "l'évitement des fautes", et s'appuie sur des techniques de conformité, la seconde permet la "tolérance aux fautes".

4.3 Fiabilité d'un logiciel et modifications

Un logiciel, que ce soit au cours de son développement ou durant son exploitation, subit des modifications. Ces modifications peuvent être classées en deux catégories : modifications correctives et modifications innovatrices. Les modifications correctives ont pour but de corriger des défauts reconnus. Si tout est mis en oeuvre pour que cette correction n'entraîne pas l'introduction de nouveaux défauts, on peut alors faire l'hypothèse que les modifications correctives accroissent la fiabilité du logiciel. Ceci ne peut cependant être affirmé que pour autant que l'on dispose de moyens techniques destinés à retester un logiciel après correction, qui permettent de s'assurer que le niveau de satisfaction atteint est égal à celui d'avant correction. Il est donc nécessaire de mettre en place des stratégies et des moyens automatisés de test qui, en l'état actuel de l'art, sont les seules garanties possibles d'une telle stabilité. Les modifications innovatrices au contraire, n'ont pas pour but la correction de défauts constatés, mais la modification des spécifications externes du logiciel. L'ampleur et la fréquence de ces modifications peuvent atteindre des niveaux excessifs qui interdisent au réalisateur le contrôle de la convergence du produit. Dans tous les cas, il est indispensable qu'une procédure formalisée et rigoureuse d'intégration des modifications soit suivie, afin de minimiser l'introduction de nouveaux défauts.

4.4 Approches qualitatives de la fiabilité du logiciel

Les approches qualitatives tendent à construire la fiabilité du logiciel. Pour ce faire, elles cherchent à atteindre deux objectifs : minimiser le nombre de défauts résiduels et minimiser les effets des défauts résiduels. Les techniques de conformité, qui ont pour but de minimiser le nombre des défauts, sont également appelées techniques d'évitement des fautes. Elles se décomposent en techniques de prévention et techniques de vérification et validation. Les techniques de prévention regroupent les structures, méthodes et outils d'aide à la définition, à la conception et au codage. Les techniques de vérification et validation ont été décrites ci-dessus. Les techniques de tolérance aux fautes tendent à minimiser les effets des défauts. Elles se décomposent également en techniques de prévention et de vérification. La prévention peut être effectuée en utilisant les méthodes d'analyse des effets des erreurs de logiciel (A.E.E.L.), analogues aux A.M.D.E.C. (analyse des modes de défaillances, de leurs effets et leur criticité) en matériel. Les techniques de vérification reposent sur la mise en oeuvre de mécanismes permettant en cours d'exécution de détecter et localiser une erreur, puis de la recouvrir par confinement, correction ou acceptation.

4.5 Approches quantitatives de la fiabilité du logiciel [BOE 75, MCC 77]

Si les techniques qualitatives permettent l'amélioration de la fiabilité, elles n'en permettent pas l'évaluation ou la prévision. Il va de soi qu'une approche contractuelle de la fiabilité passe non seulement par l'assurance de la qualité des travaux c'est-à-dire la maîtrise des techniques qualitatives, mais aussi par la connaissance chiffrée de la qualité effectivement observable sur les produits. Ce second type d'approche reste encore, il faut le dire, un domaine de recherche fondamentale et appliquée. Il n'en reste pas moins que de nombreux travaux ont été effectués sur ce thème et que certains semblent être proches de l'applicabilité industrielle. Les approches quantitatives peuvent être scindées entre approche prédictive et approche après codage. L'approche prédictive s'appuie sur des résultats de mesure prélevés soit sur la documentation (Mc CALL, BOEHM), soit sur le code (id, + Mc CABE et HALSTEAD). Ces mesures, fournies à un modèle prédictif, sont susceptibles de donner une indication sur la fiabilité du produit final. Si l'approche peut sembler discutable du point de vue fiabiliste, elle permet en revanche de connaître certaines propriétés des diverses descriptions du logiciel (des spécifications externes au code), propriétés qui reflètent les problèmes qui seront rencontrés.

L'approche après codage se répartit également selon deux orientations. la première repose sur des modèles d'estimation de la fiabilité, alimentés en données prélevées durant les tests ; ces données peuvent être relatives au bon fonctionnement observé (modèles de MUSA, LITTLEWOOD-VERRAL, GOEL-OKUMOTO, ...) ou au taux de couverture des tests (NELSON, MILLS...).

La seconde consiste à mesurer des intervalles de défaillance sur le produit réputé opérationnel (éventuellement en service effectif) et à fournir ces données à un modèle d'évaluation de la probabilité de bon fonctionnement (modèle de MUSA, en particulier).

L'ensemble de ces approches qualitatives et quantitatives de la fiabilité du logiciel est résumé sur la figure 1.

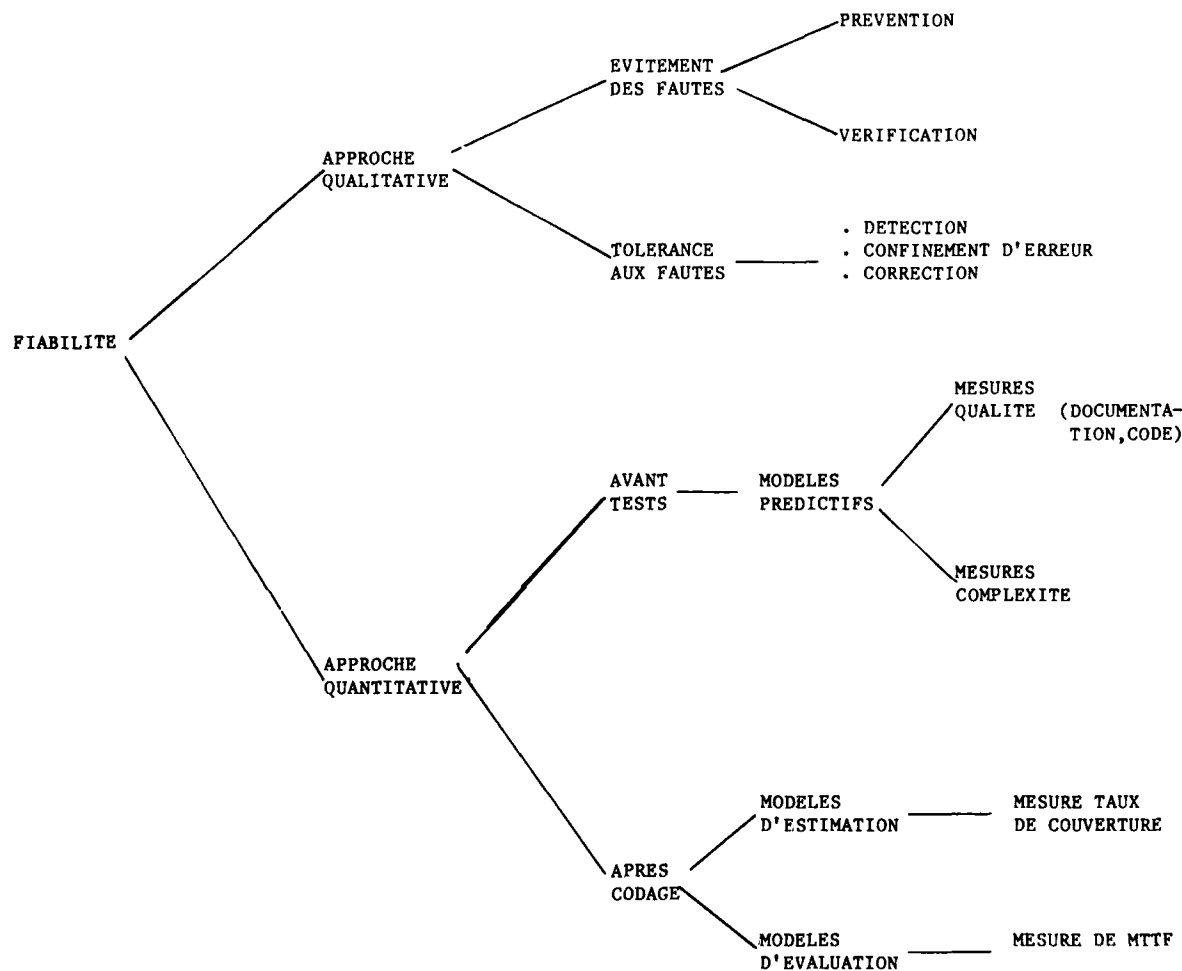


FIGURE 1 : APPROCHES ET TECHNIQUES EN FIABILITE DU LOGICIEL

5. LE DEVELOPPEMENT DES LOGICIELS MIRAGE 2000

Les logiciels centraux des divers avions Mirage 2000 représentent un exemple d'effort industriel pour assurer un degré élevé de fiabilité d'un produit dont le rôle est essentiel au bon fonctionnement du système de navigation et d'armement. Ces logiciels constituent une famille dont l'architecture et les mécanismes généraux sont communs, mais qui présente des variantes importantes selon la nature des missions et des équipements affectés à un type d'avion. De façon à assurer la qualité des logiciels produits, tout en respectant les délais très courts et les budgets établis, l'ESD a mis en place une méthodologie, MINERVE, et des outils qui assurent son support. Ceci ne constitue que le début d'une politique visant à mieux maîtriser le processus de production de logiciel, politique qui est prolongée par des études en génie logiciel présentées plus loin.

5.1 La méthodologie MINERVE (formalisée depuis 1976) [PER 79]

Elle repose sur trois principes :

- les travaux sont découpés en phases et en étapes caractérisées par des activités, des produits et des responsabilités clairement définis,
- La qualité des produits ainsi que leurs coûts et délais de réalisation sont contrôlés de façon continue,
- Les modifications sont prises en compte quel que soit le degré d'avancement des travaux, selon une procédure unique destinée à éviter toute dégradation de la qualité du logiciel.

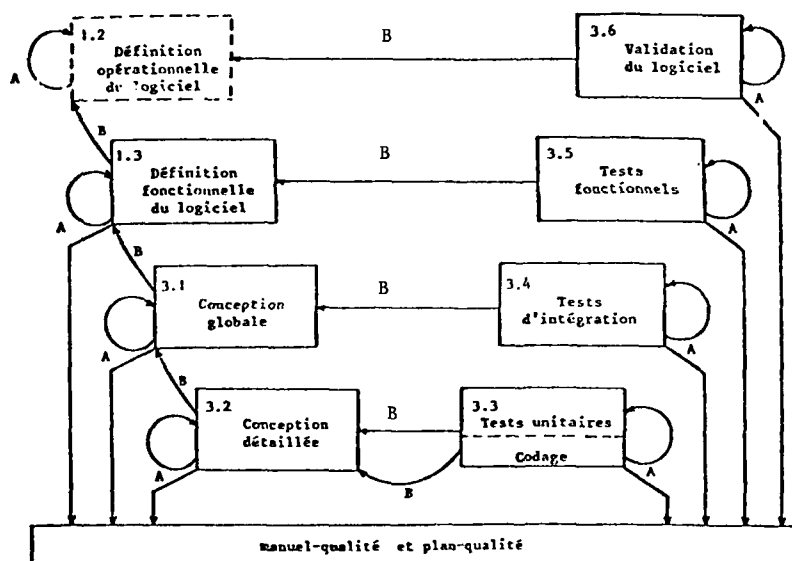
MINERVE constitue le manuel-qualité du logiciel de l'ESD.

Pour chaque projet, des dispositions spécifiques sont établies dans le cadre de MINERVE et en accord avec le groupe "qualité du logiciel" ; elles constituent le plan-qualité du projet.

MINERVE distingue 2 types de contrôles :

- Contrôles de type A : ce sont des contrôles internes au produit d'une étape ; effectués par relecture et analyse des documents ou du code, ils consistent à vérifier que le produit respecte les règles précisées dans le plan-qualité (en particulier l'application des standards de documentation, de codage, d'utilisation d'outils...) ; on s'assure également de la cohérence interne, de la complétude, de la lisibilité et de la précision de chaque produit.
- Contrôles de type B : ce sont des contrôles de cohérence entre un niveau de description du logiciel et le niveau antérieur (code vers la conception, conception vers la définition) ; ce type de contrôle, comme le précédent, est réalisé sous forme de relectures et de revues. En outre, ces contrôles incluent les tests des programmes qui constituent actuellement le seul moyen de vérifier que le comportement effectif du logiciel est conforme à la description qui en a été faite. Ce dernier mode de contrôle peut également être réalisé par l'utilisation de maquettes ou d'analyseurs de spécifications.

Le passage d'une étape à l'autre est conditionné par l'obtention de résultats satisfaisants pour ces deux types de contrôle. L'ensemble des contrôles est schématisé par la figure 2.



La gestion des modifications dans MINERVE constitue une originalité : de par sa procédure, elle intègre la maintenance dans le même processus que le développement.

La procédure de modification se déroule en deux temps :

- Prise en compte de la modification : un constat d'anomalie, ou une demande d'évolution entraîne la rédaction d'une "fiche de modification", qui doit être suffisamment précise pour que l'ensemble des produits mis en cause (code, conception, définition) soit déterminé et que l'impact sur les coûts et délais puisse être évalué.
- Réalisation de la modification : décidée en connaissance de cause, la modification peut alors être réalisée en respectant la séquence des étapes concernées, ce qui garantit la cohérence de l'ensemble des produits ; l'exécution de la modification est consignée sur une "fiche suiveuse".

5.2 L'outillage utilisé pour le Mirage 2000

L'application rigoureuse d'une méthodologie permet d'obtenir une meilleure qualité des produits et une meilleure maîtrise des coûts et délais. Elle permet également de diminuer les coûts d'intégration et de mise au point finale, ainsi que ceux de maintenance ultérieure. En revanche, elle met en évidence des travaux bien formalisés, qui se prêtent dès lors à une automatisation. Ici comme dans d'autres domaines d'activité, l'automatisation vise à augmenter la productivité, à réduire les tâches fastidieuses et répétitives et à faciliter l'obtention d'une qualité stable.

L'ensemble des documents produits est géré sur des machines de traitement de textes, qui sont interconnectées avec l'ordinateur central utilisé pour la production de programmes. Ainsi, les ingénieurs peuvent-ils consulter et mettre à jour cette documentation directement depuis leur terminal d'ordinateur, tandis que le personnel de secrétariat dispose d'un matériel adapté à la saisie massive. En ce qui concerne les spécifications fonctionnelles, des diagrammes de décomposition permettent d'en rendre l'appréhension beaucoup plus aisée, en particulier en matière de description de la hiérarchie des fonctions et de leurs interdépendances.

Pour la conception du logiciel, l'ESD a développé un "langage de description de modules" [CAT 83] qui permet d'exprimer la structure globale des programmes et de décrire les interfaces de modules. Ceci complète le langage LTR-V2, utilisé pour l'essentiel du codage, dans le sens de l'évolution actuelle des langages à compilation séparée (LTR-V3, ADA...). Ce langage est analysé par un préprocesseur qui, si la description fournie est correcte, engendre le code LTR-V2 correspondant.

En ce qui concerne les tests, deux types d'outils ont été développés :

- Le langage LOTUS, qui permet l'écriture symbolique de programmes de test manipulant les objets du programme à tester sous leurs noms et types LTR. LOTUS permet ainsi la formalisation tant des procédures que des résultats de test.
 - Les baies de validation de logiciel (B.V.L.) [BOU 79] qui permettent la simulation en temps réel de l'environnement aux bornes du calculateur opérationnel, l'analyse automatique des réponses du logiciel aux stimuli et la possibilité pour un opérateur de surveiller l'exécution et d'engendrer des événements à tout instant.
- Ces deux outils de test, qui permettent l'accès aux programmes opérationnels en symbolique et avec les formats externes des données, autorisent également le test automatique, de par leur capacité à enregistrer des scénarii stimuli-réponses et à les restituer de façon autonome. De tels moyens, bien qu'incomplets au regard des problèmes liés à la fiabilité du logiciel, nous paraissent indispensables.

Enfin, les fiches de modification sont également informatisées, afin de faciliter leur consultation, leur suivi et leur diffusion.

6. LES EFFORTS DE L'ESD EN MATIERE D'ASSURANCE-QUALITE

L'exemple concret du Mirage 2000 présenté ci-dessus est bien représentatif des efforts menés par l'ESD en matière d'assurance-qualité du logiciel. Il est également important de noter que l'ensemble des procédures de construction de la qualité est mené par les équipes de réalisation. Un "groupe qualité du logiciel" intervient en support, pour aider à la rédaction des plans-qualité, à la mise en oeuvre d'outils et à certaines procédures critiques d'évaluation. Mais la responsabilité des réalisateurs vis-à-vis de la qualité des produits est érigée en principe directeur, ce qui constitue un moteur puissant d'intégration de la qualité dans le processus de développement. Ainsi, le respect des règles assurant la qualité n'apparaît-il pas comme une contrainte extérieure, et quelque peu antagoniste avec les objectifs opérationnels.

Afin d'accroître sa maîtrise de la qualité du logiciel, à la fois dans la perspective de logiciels de plus en plus volumineux et dans le but d'effectuer des gains de productivité, l'ESD a conduit de nombreuses études en génie logiciel. Trois d'entre elles constituent la base de l'effort en cours.

6.1 D.L.A.O. : Définition de Logiciel Assistée par Ordinateur [CHE 82]

DLAO est un langage et un ensemble d'outils qui constituent un système d'aide à la définition du logiciel. Cet effort est soutenu par la DRET. Le langage DLAO est basé sur un modèle entité-relation. Il distingue un certain nombre d'objets, possédant des attributs et entrant dans des relations-types. Si le langage permet d'exprimer le squelette essentiel de la spécification, il autorise aussi le recours à des formulations en langage naturel qui permettent d'explicitier certains aspects sémantiques. Il s'agit donc d'un langage semi-formel, lisible par un non-spécialiste. Divers outils interactifs permettent la rédaction, la consultation, la modification, l'archivage et la réutilisation des constituants d'une spécification.

En particulier, à partir d'une spécification d'origine, le système peut constituer des listes de références croisées, de flux de données, des arbres exprimant les structures de contrôle, des graphes de transition etc...

6.2 IDA : langage de test et outils associés [LAM 82]

De l'expérience des B.V.L. est née l'idée d'un langage de test de logiciels temps réel, adaptable à divers calculateurs et langages de programmation. En effet, les BVL constituent un système spécifique aux calculateurs ESD et au langage LTR.

Le système IDA réside sur un ordinateur, dit "machine de test", relié au ordinateur opérationnel par une interface qui permet d'effectuer des opérations de contrôle : lire, écrire, lancer, arrêter..., et de détecter des événements : exécution d'une instruction précise, changement d'état d'une donnée... Le langage permet d'écrire des programmes de test qui s'exécutent sur la machine de test et manipulent le programme opérationnel sous test en utilisant ses symboles, sans se préoccuper de la gestion de l'interface entre calculateurs.

Il est également possible de créer des outils de test, invocables depuis le langage et qui en constituent des "fonctions standard". Un tel outil a déjà été développé à l'ESD pour contrôler un modèle de comportement de logiciel exprimé sous forme de réseaux de Pétri.

Un sous-ensemble restreint d'IDA (hors temps réel) est opérationnel dans le cadre des projets Mirage 2000 : LOTUS. Un autre sous-ensemble, fonctionnant en temps réel cette fois est un développement.

6.3 Evaluation de la qualité du logiciel [BOE 75, MCC 77]

Si les deux projets décrits ci-dessus constituent des efforts à caractère principalement qualitatif, ils ouvrent également des perspectives sur la quantification de la qualité du logiciel. DLAO permet d'effectuer des mesures de complétude, de cohérence et de complexité sur la définition. IDA permet d'effectuer des statistiques sur les tests, et en particulier permettrait, en développant les outils correspondants, la mesure du taux de couverture.

D'autre part, deux études sont en cours, afin d'établir une terminologie et des règles d'évaluation de la qualité et de la fiabilité des logiciels. Ces études sont soutenues par la DTEN.

L'évaluation de la qualité s'appuie sur les travaux de BOEHM et de Mc CALL. Elle s'effectue à partir d'un modèle qui définit des facteurs (vision du client) de qualité, qui s'apprécient à partir de critères (vision interne du réalisateur). Les critères sont eux-mêmes évalués à partir de métriques, éléments quantifiables de base. Cette approche de la qualité du logiciel permettra de définir des "classes de qualité", qui pourront donner lieu à engagement contractuel.

L'état de l'art sur les problèmes de fiabilité ayant été résumé plus haut, cette autre étude n'est pas développée ici.

A l'ESD, les retombées de ces études vont résulter dans la mise en place d'expérimentations s'appuyant sur les moyens disponibles, sachant que toutes les mesures ne peuvent être réalisées aujourd'hui (sinon à la main, ce qui est économiquement dissuasif). Ces expérimentations permettront d'apprécier l'intérêt et la faisabilité industriels de ces approches quantifiées.

7. CONCLUSION

La certification de logiciel n'est pas encore réalisable au plan industriel. Le présent article a cependant mis en évidence le fait que des techniques et des outils ont atteint un stade qui permet d'expérimenter des procédures de certification ; en effet, pour ce qui est des approches quantitatives, seules des bases statistiques et des expérimentations permettront de valider les procédures.

La démarche proposée, qui consiste à construire la qualité, en particulier la fiabilité, d'un logiciel au long de son développement, et en parallèle à mesurer les produits résultants, nécessite qu'un ensemble intégré de méthodes et d'outils soit mis en place, afin d'assurer la cohérence et la continuité des activités. C'est vers un tel ensemble que s'oriente l'ESD, dans le cadre de la méthodologie MINERVE. Les divers moyens de développement (ordinateur de production de programmes, machines à traitement de textes, diverses stations de test...) sont interconnectés afin de constituer un atelier intégré de génie logiciel. Au cœur de cet atelier sera constituée à terme une base de données unique, banalisant l'accès à toutes les informations pour les divers outils.

En résumé, les perspectives en matière de certification de logiciel peuvent être envisagées comme étant la mise au point de procédures permettant de s'assurer d'un "degré de fiabilité". Pour que ces procédures puissent s'appliquer, il est nécessaire que la formalisation des activités de développement (spécification, tests...) soit augmentée. Dans un contexte industriel, ces procédures devront naturellement s'intégrer aux opérations de vérification et de validation, qui offrent la structure technique et administrative nécessaire. Ce besoin d'une visibilité accrue tant des activités que des produits signifie que la certification du logiciel ne pourra être réalisée qu'en s'appuyant sur la "validation du processus de production".

BIBLIOGRAPHIE (Abrégée)

- [BOE 75] B.W. BOEHM
"Characteristics of software quality"
TRW - North Holland - 1975
- [BOU 79] E. BOUTHORS
"Simulateur pour intégration avec tests dynamiques des logiciels de calculateurs centraux embarqués"
AGARD - Septembre 1979
- [CAT 83] P. CATEL
"Vers une modularité du logiciel conçue pour les besoins de l'utilisateur"
AGARD - Ottawa - Avril 1983
- [CHE 82] S. CHENUT - MARTIN et F. DOLADILLE
"DLAO : un système d'aide à la définition de logiciels avioniques"
AGARD - La Haye - Septembre 1982
- [DO 178] RTCA - EUROCAE
Software considerations in airborne systems and equipment certification
RTCA DO-178/EUROCAE Ed. 12, May 1982
- [HOW 82] W. HOWDEN
"Life -cycle software validation"
IEEE Transactions on S.E. 1982
- [LAM 82] G. LAMARCHE et P. TAILLIBERT
"IDA : Langage de test et outils associés"
AGARD - La Haye - Septembre 1982
- [LAW 81] R.W. LAWLER
"System Perspective on software quality"
COMPSAC'81
- [MCC 77] J.A. Mc CALL et al.
"Factors in software quality"
GEC 77 CIS02 - Juin 1977
- [PER 79] J. PERIN
"Logiciel avionique : expérience pratique d'une méthodologie"
AGARD - Ottawa - Avril 1979
- [RGA 82] RG Aéro 00026
"Construction et assurance de la qualité en conception"
Bureau de la Normalisation de l'Aéronautique et de l'Espace
Groupe de travail GT74 - Septembre 1982
- [SOR 79] A. SORKOWITZ
"Certification testing : a procedure to impose the quality of software testing"
COMPUTER - 1979
- [TRO 79] R. TROY
"Basic concepts in software for reliable computing"
SURF - Septembre 1979

AD P 003 400

VERIFICATION AND VALIDATION TECHNIQUES APPLIED TO THE RECONFIGURABLE BLOCK 5D SATELLITE SOFTWARE*

by
William C. Diniak
The Charles Stark Draper Laboratory, Inc.
555 Technology Square
Cambridge, Massachusetts 02139
U.S.A.

SUMMARY

The U.S. Air Force Defense Meteorological Satellite Program (DMSP) Block 5D satellite uses two general purpose digital flight computers for control of the satellite and launch vehicle. During the life of six satellites, various reconfigurations of the flight software were accomplished to achieve improvements or to compensate for problems. An independent verification and validation (IV&V) process was conducted by CSDL on that flight software prior to launch and prior to uplink of major software changes. To accomplish this, CSDL developed and used a closed-loop digital simulation of the mission to exercise the actual flight software in a realistic environment. This paper describes this experience, focusing on how both the simulation and the IV&V process evolved as a consequence of the flight history.

DMSP BLOCK 5D CONCEPT

The primary mission of the U.S. Air Force DMSP is to gather weather data for military use. Conceptually, the data are obtained from a variety of sensors carried by 3-axis stabilized Block 5D satellites circling the earth at an altitude of 450 nautical miles (833 km) in near-polar sun synchronous orbits. With this orbit, full coverage of the earth is made by each satellite in 12 hours, while scanning an 1800-nmi (3330 km) wide area.

The primary meteorological payload is a Westinghouse Operational Line Scan System (OLS) that takes visual and infrared imagery for use in analyzing cloud patterns. The OLS also manages the data from a group of other meteorological sensors which can measure parameters such as temperature, moisture, ion and electron densities, and x-ray emissions.

BLOCK 5D DESCRIPTION

The Block 5D satellite is manufactured by the Government Systems/Astro-Electronics Division of the RCA Corporation. The satellite is structurally divided into the following four major sections. (See Figure 1.)

- (1) A rigid, precision mounting platform holds those sensors which require precise alignment relative to the OLS and pointing toward the earth. Among this equipment are:
 - (a) One inertial measurement unit (IMU) containing:
 - (1) Three orthogonal, strapdown gyros for measurement of angular velocity during launch and in orbit. A fourth, skewed gyro provides redundancy in orbit when selected for use by the flight software.
 - (2) Three orthogonal accelerometers measure specific force during launch for navigation of the launch vehicle.
 - (b) One strapdown, six-slit celestial sensor assembly (CSA), which indicates star crossing events for use in precision attitude determination.
 - (c) One earth sensor assembly (ESA), containing four quadrants which view the carbon dioxide horizon to measure pitch and roll errors for backup attitude determination.
 - (d) One sun sensor unit (SSU) to measure yaw error for backup attitude determination. The yaw error is also estimated from the roll gyro output while in a yaw-gyrocompass submode.
- (2) An equipment support module is a central structural section which encloses much of the satellite electronics including:
 - (a) Three orthogonal reaction wheel assemblies (RWAs) which provide attitude control torque. A fourth, skewed reaction wheel provides redundancy when selected for use by the flight software.
 - (b) Two sets, for redundancy, of dual orthogonal magnetic momentum unloading (MMU) coils to provide low-level torques through interaction with the earth's magnetic field.
 - (c) Two SCP234 16-bit CMOS digital computers for control of the mission.
 - (d) One controls interface unit (CIU), a switching center which interfaces the flight computers with the various satellite equipment.
- (3) A reaction control equipment support structure encloses the spent upper stage Apogee Kick Motor (AKM) and also contains:
 - (a) Eight nitrogen (N_2) thrusters which provide low-level attitude control torque during launch.

*The activities described in this paper occurred during work performed by CSDL for the USAF Space Division under contracts F04701-74-C-0020 and F04701-78-C-0018.

- (b) Four hydrazine (N_2H_4) thrusters which are used for final velocity trim and, along with the nitrogen thrusters, provide high-level attitude control torque during launch.
- (4) A solar array is rotationally driven about the satellite pitch axis to point the array normal toward the sun. The array of solar cells generates electrical power for distribution to satellite equipment, including batteries, through a direct energy transfer system.

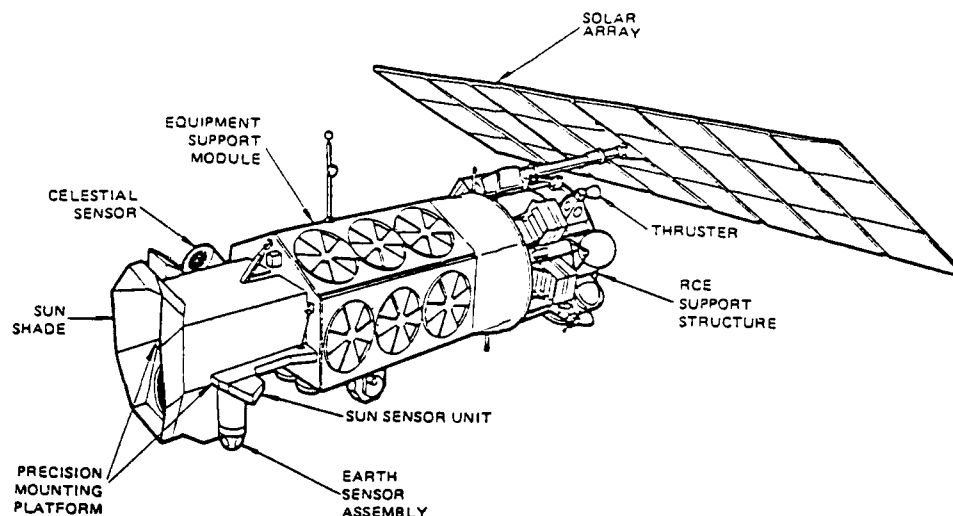


Figure 1. The Block 5D-2 satellite.

The satellite contains S-band communication equipment for the receipt of command messages and computer memory words uplinked from the ground to each of the flight computers or CIU. With this feature, satellite ephemeris information is periodically refreshed, or the flight software can be reconfigured as desired.

Provision is also made for the storage of meteorological data on tape recorders and the S-band downlink of both meteorological sensor data and housekeeping telemetry data from both flight computers and other equipment.

The first generation of Block 5D satellites, called 5D-1 and comprising F1 through F5, were launched using McDonnell-Douglas (MDAC) Thor LV-2F boosters and two upper stages of solid propellant rocket motors. Ascent guidance, navigation, and control of the launch vehicle was completely controlled by MDAC software within one of the SCP234 computers. The launch dates and life-spans of these vehicles as of September 1983 are shown below.

Satellite	Launch Date	Life (months)
F1	12 Sept. 76	36
F2	4 June 77	32
F3	1 May 78	65
F4	6 June 79	14
F5	15 July 80	0
F6	20 Dec. 82	9+

The current generation, called 5D-2, is somewhat longer and heavier, but with similar satellite equipment and appearance, and is launched by a General Dynamics Atlas F Booster and one upper stage Apogee kick motor. Ascent guidance, navigation, and control (GN&C) through Atlas booster propulsion is provided by a ground-based radio guidance system. Beyond the booster portion of flight, MDAC software within an SCP234 computer provides full guidance and control of the launch vehicle.

After completion of propulsion and deployment, software control is handed over from MDAC ascent to RCA orbital programs within the SCP234 flight computers. The orbit software is a mixture of Honeywell and RCA programs. Although the original F1 flight loads were operable, improvements were incorporated throughout the 5D-1 era and, for 5D-2, expansion of the SCP234 memory capacity from 16 to 28 kilobytes permitted considerable additional improvements.

Generally, each of the two flight computers contains identical orbit software programs and is fully capable of controlling the satellite attitude, managing angular momentum, and pointing the solar array. Normally, both computers are always running, have sequential access to equipment data, and both routinely attempt to command the various external satellite devices. However, they are each in different modes of operation and only one computer is actually connected to a device (such as the reaction wheels) at a given time via the controls interface unit (CIU). In the event of unsatisfactory performance, monitoring functions within the CIU or within that SCP234 designated as the master can cause automatic changes in the modes of operation.

With the flight computers playing such a vital role in the launch vehicle and satellite systems, it is important that the software perform correctly. USAF Space Division policy stipulates that the integrity of the flight software be demonstrated and certified prior to flight by a technical organization independent of the system developers. The focus of this paper will be on CSDL's activities toward providing independent verification and validation of the flight software.

SOFTWARE V&V CONCEPTS

It will be helpful to first state some definitions of the words "verification" and "validation" as used with regard to software. The rudiments of a system development process are indicated by Figure 2, along with a suggestion of the V&V components associated with it. In a modern development process, the mission or user requirements are documented in a system specification so that the design and test goals are defined. The system specification must clearly state the system requirements and will generally partition the requirements into those which shall be met by either software or hardware. Software requirements form the basis for a detailed design, as documented in a software design specification. The design is then implemented by coding into the operational software.

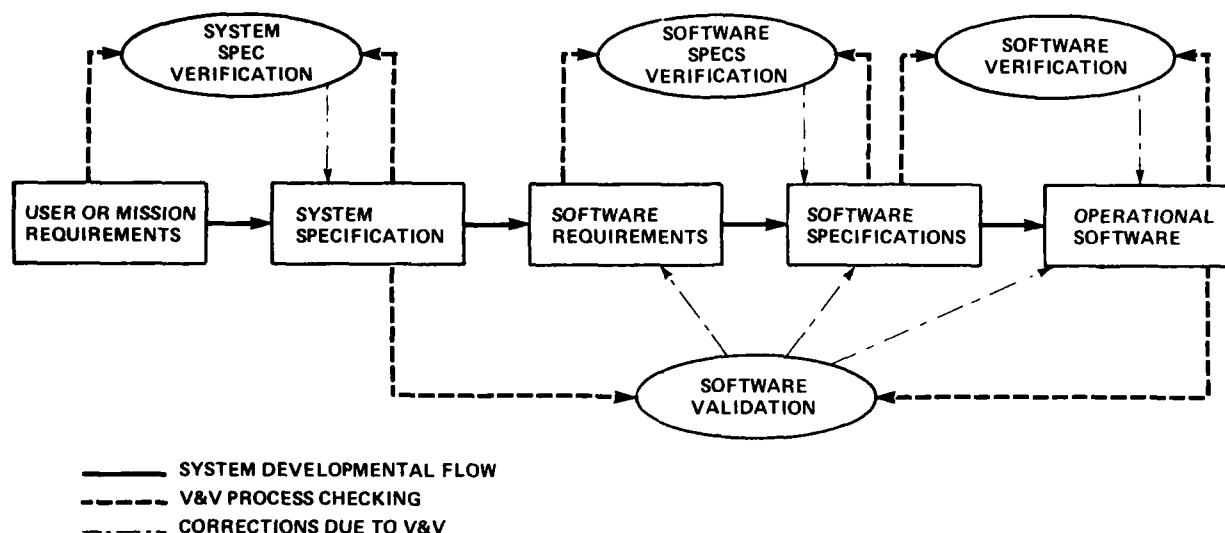


Figure 2. Product development process.

Verification is that process of confirming the traceability and accurate implementation of the final product through the successive stages of specifications and procedures which govern the development cycle. Software verification determines whether a computer program was developed in accordance with the software specification. It will not necessarily guarantee that the software is usable, only that it satisfies its specifications.

Software validation refers to the process of determining that the actual software will satisfy the mission requirements, that it will perform the necessary functions in the mission environment.

The most recent CSDL IV&V cycle for Block 5D was on F6, the first of the 5D-2 vehicles. The CSDL activities, indicated in detail in Tables 1 and 2, included painstaking analysis of each line of code and realistic exercise of all flight code functions on the Multiprocessor Software Test Facility (MSTF) closed-loop simulation.

MULTIPROCESSOR SOFTWARE TEST FACILITY

The MSTF is a closed-loop digital test bed for the execution of actual, unmodified Block 5D flight code in a wide variety of scenarios which span the mission from prelaunch into orbit. In general, runs proceed automatically from initial conditions specified by the data base, modified as desired by the operator. Almost any type of simulated hardware failure condition can be introduced by the operator during a run.

The MSTF hardware configuration, as shown in Figure 3, includes two RCA SCP234/GT computers containing the flight code, and a PE8/32D host minicomputer containing models of all flight vehicle equipment which interact with the flight computers. Block diagrams of the MSTF indicating the ascent and orbit simulation modules are shown by Figures 4 and 5.

The MSTF was developed by CSDL to replace the Software Test Facility (STF) used for Block 5D-1 software validation. The development included minimally modifying 5D-1 hardware models as required for 5D-2, interfacing the host computer with RCA-developed ground test versions of the SCP234s and CIU, and completely redesigning the executive control software. Table 3 shows some of the significant differences between the STF and the MSTF and provides information on the character of the simulation. Because the MSTF and, in fact, the entire V&V process evolved from an earlier CSDL effort on 5D-1, there is benefit in a retrospective review of that history.

CSDL SIMULATION DEVELOPMENT EFFORT

The CSDL Block 5D activity for USAF was initiated in late 1973 as part of an Independent Dynamic Verification Program to satisfy the SAMSO Commanders Policy as specified in SAMSO pamphlet 800-4. Such a program requires the following fundamental tasks:

- (1) Model development and verification.
- (2) Simulation development and dynamic performance verification by simulation.
- (3) Supporting analyses.

Table 1. CSDL verification activity on Block 5D F6.

Current Verification Techniques	Benefits
Maintain permanent files of all flight code images delivered to CSDL in central computer facility	Permits automatic comparison of instructions and parameters against other files. Permits searching for shared memory locations
Create central computer data base files from information contained in data base specifications (such as numeric value, units, scaling, range, address)	Permits automatic comparison of specified data against actual flight loads, and consistency checking of specified information
Transfer flight code images from simulation tests into central computer files	Comparisons against delivered flight loads show any modified instructions and parameters
Automatic file comparison, on a module-by-module basis, of all flight code deliveries against previous versions. Differences are individually reconciled, and audit reports issued to government within a week	Auditing of deliveries permits flight load configuration to be tracked and correlated against documentation change notices
Flight code assembly listings are manually analyzed, line-by-line, and the logic is compared against pertinent specifications. Results reported in a software verification memo (SVM) for each module	Analysis uncovers errors in both coding and specifications. Documentation by SVM results in action toward convergence of specifications and coding. This verification effort enhances familiarity of software and assists in simulation test planning
Flight code analysis includes the generation of a highly detailed, as-coded flow diagram for a module, which is included in the SVM	The flow diagrams are useful to the DMSP community, and demonstrate the level of the analyst's effort (at the expense of slowing down the analysis process). The diagrams show off-nominal paths which should be tested
The verification flow diagrams show the binary point scaling at every arithmetic operation	Forces analyst to check for consistency of scaling and shows potential for overflows at intermediate operations
Software verification notebooks (SVNs) are compiled for each flight code module containing all specification, memo, SVM, and file compare information relevant to the module	The SVN for each module, which shows its complete history, is useful to analysts as a quick comprehensive reference
Automatic file comparison tool used to compare portions of modified modules against previous versions	Reverification of modules, required by flight code or specification changes, is expedited by permitting some automatic analysis of changes

Table 2. CSDL validation activity on Block 5D F6.

Current Validation Techniques	Benefits
Unmodified flight code, imbedded in the MSTF closed-loop simulation, is exercised in a variety of mission scenarios	Operation of actual code demonstrated in a realistic environment produces maximum confidence that the software is operable. Problems with the flight code automatically manifest themselves by peculiar test results independent of the intention of the test planner
For each test, a standard network of flight computer breakpoints is set in off-nominal logic paths	Permits accounting for paths exercised. Gives analyst insight into conduct of test, including confirmation that test scenario is achieving intended results
All telemetry data issued by flight computers from each test are stored and automatically post processed to produce plots and formatted dumps	Plots permit analyst to at least quick scan all data to detect anomalous results which may not be related to the prime test objective
Internal test reviews are held to consider the results of each test. Formal data reviews are later held for the government and associate contractors	Improves the analysis of data and promotes understanding of flight code behavior
A comprehensive test analysis report is written 30 days after the conclusion of testing	Test results are widely distributed throughout the DMSP community to promote understanding of the launch vehicle and satellite performance

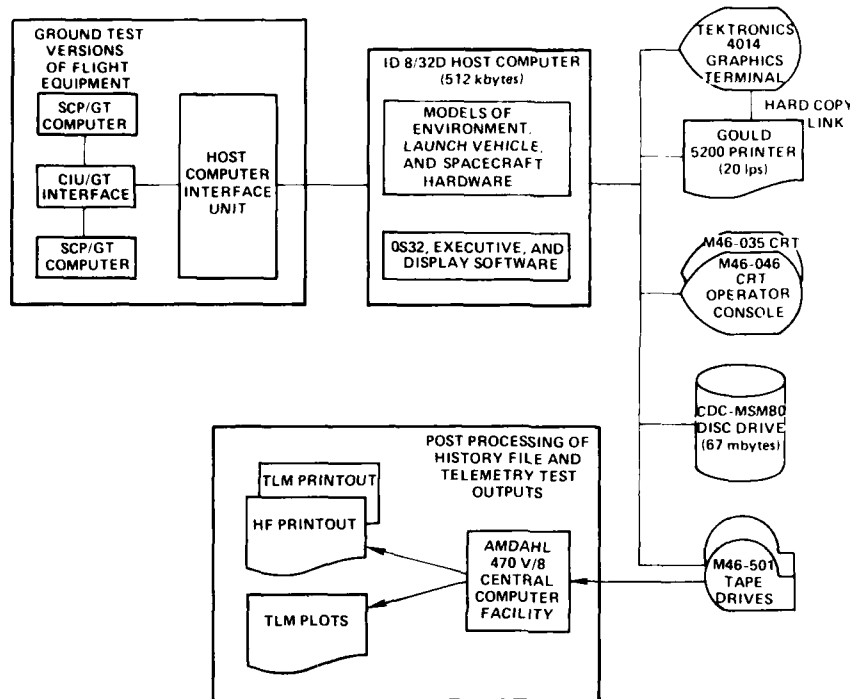


Figure 3. The multiprocessor software test facility.

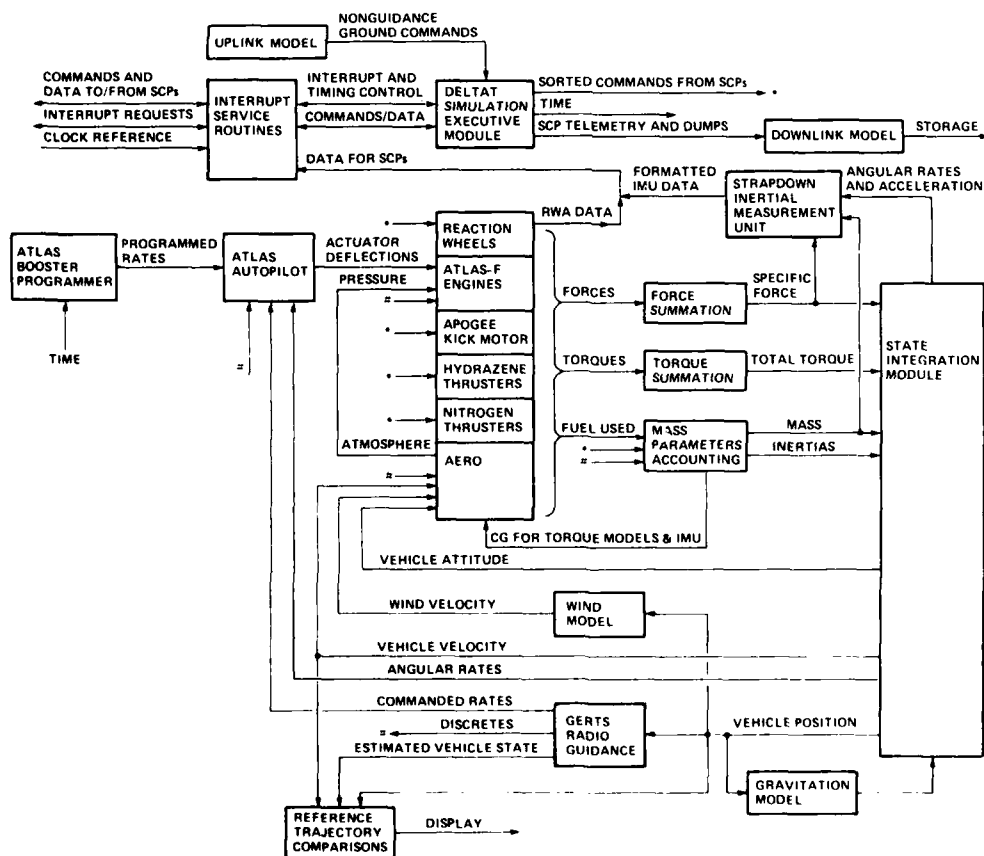


Figure 4. MSTF prelaunch and ascent modules.

Table 3. Comparison of CSDL STF and MSTF simulations.

STF (for 5D-1)	MSTF (for 5D-1 and 5D-2)
Single, flight-worthy RCA SCP234 computer with 16 kbytes of memory	Two ground test versions of 28 kbyte RCA SCP234, made with commercial-quality parts
Primitive software simulation of the controls interface unit	Ground test, hardware version of portions of the flight CIU. Contains specialized serial input and output buffers
MIL computer interface device containing fundamental timing source for host and flight computers	Ground test interface unit containing crystal oscillator timing source for flight computers
Host computer was ID70 with 64 kbytes memory, later upgraded to PE 8/32C computer with 128 kbytes memory	Host computer is PE 8/32D minicomputer with 512 kbytes memory
Sequential operation of SCP234 and ID70 computers, limiting operating time to 10-20 times slower than real time	Parallel operation of SCP234 and PE 8/32D computers, but forced to stay within Δt of each other. As low as 4 times slower than real time
Home-grown executive and cannibalized disc operating system with limited features	Perkin-Elmer OS32 operating system software with wide variety of command features
No run time graphics	Run time graphics of up to 20 multivariable plots with capability of hard copy onto line printer
No capability for scheduling automatic operator commands	Capabilities for time-scheduling operator commands and for issuing catalogued sets of operator commands
No breakpoint capability for SCP234	Capability to breakpoint all or any locations in both SCP234/GT computers
Checkpoints of simulation contents limited to 28 by 9-megabyte disc system	Number of checkpoints greatly increased by use of 67-megabyte disc
No routine storage of simulation data, except SCP234 telemetry output, on magtape	Capability for history file creation on mag tape to record simulation data and run events
Contained feature to command detailed printout of all I/O traffic	May command all, or only operator-defined subset of I/O traffic to conserve running time
Original modeling caused interrupts to SCP234 from IMU, ESA, and downlink model to be synchronous with model update times	For more faithful representation, simulation model interrupts are scheduled asynchronously with model update times

The role of the minicomputer was to simulate the relevant hardware and physical environment outside the flight computers with sufficient fidelity as to routinely respond to SCP commands, issue satellite equipment interrupts to the SCP, and provide meaningful formatted input data to the SCPs when requested.

Accomplishing this within 64 kbytes of memory, absurdly small by today's standards, proved to be quite a challenge, but was successfully accomplished by coding in assembly language, using overlay techniques, minimizing the resident operating system and repertoire of operator commands.

The scarcity of SCP234 flight computers resulted in the inclusion of only one SCP234 in the STF simulation, whereas the satellite actually contains two. Testing in a single computer environment was nervously considered tolerable because the flight computers were not planned to directly interact, and would contain identical orbit flight code. Some thought was given in the preliminary simulation design to simulating a dual computer situation by sequential storing and loading the contents of dual computers to/from disc storage during simulation operation. However, the complexities in preserving timing fidelity and a terrible penalty on simulation run time, reinforced the decision to make do with a single SCP234. As mechanized, the STF simulation ran approximately 10 to 20 times slower than real time under ordinary modes of operation.

A comparison of STF and MSTF characteristics is indicated by Table 3. Other STF and MSTF features of interest include

- (1) A large network of logic which produces hard copy alert messages upon failure of reasonableness or protocol tests. This exposes flight code and simulation problems which might otherwise be obscured by observing only closed-loop performance.
- (2) A checkpointing feature which permits storage on disc, on command, of the memory and register contents of the flight and host computers. Selectively storing the simulation state at a given time greatly facilitates the starting, continuation, or rerun of tests.
- (3) Provision of a reference trajectory file of state vectors for ascent which permits automatic, real-time comparison of results against the designed trajectory.

- (4) Positive response to all flight computer input/output traffic. For special use, the entire sequential I/O stream can be printed for analysis (at an agonizingly slow rate).
- (5) A generous network of flags, biases, and scale factors is included in hardware models to facilitate the setup of satellite hardware off-nominal conditions.

In the original simulation modeling effort, use was made of a diverse variety of reference information from the various hardware contractors. At that time, there did not exist a coherent body of documentation on which to base a simulation design. Each contractor was tasked to review portions of the CSDL design in order to minimize future disagreements in the resolution of simulation-indicated anomalies. This exercise tended to benefit all parties by highlighting system interface problems.

There was initial uncertainty on the level of modeling fidelity required. The absence of closed-loop simulation facilities in the DMSP program at the time struck an interest in the possible multipurpose use for the STF. Consequently, the approach was taken to model at the highest level supportable by the fundamental simulation computational time interval (0.020 second in orbit, and 0.010 second in ascent). The computation rates were selected to be higher than the planned flight software cycle rates. Generally, modeling overkill tended to be appreciated in later years.

From flight to flight during the 5D-1 era, minor changes to the STF were implemented due to (1) differences in each flight vehicle and (2) enhancements to the STF due to operational experience. (See Table 4.) However, changes were always made in a manner such that executable code would be downward compatible, so that the simulation could be ready within an hour in support of problem solutions or any of the operational Block 5D satellites.

The 5D-1 experience proved the usefulness of the STF. However, the lack of dual SCP234s and the possibility of unobservable interactions prompted the development of the MSTF currently used for either 5D-1 or 5D-2.

CSDL VALIDATION EFFORT FOR FLIGHT 1

Prior to Flight 1, the desire to conserve program resources and the state of published flight code specifications and other documentation did not justify the tasking of a formal verification process.

To achieve maximum confidence in the success of F1 in a minimum period of available test time (several months, stretching to eight as launch was delayed), CSDL effort was concentrated on a variety of ascent performance tests, augmented by several nominal runs (which exercised approximately 80% of the orbit code. Later, with the satellite in orbit, testing of the orbit code could be continued with the potential luxury of reprogramming that code if bugs were detected. The detection of ascent problems, however, could not be delayed since they might result in loss of mission.

Several problems in the ascent and orbital code were found by CSDL during this testing and resolved prior to launch. The most noteworthy was a bug in the ascent guidance software which caused a catastrophic chain of events if automatic orbit retargeting were necessitated by a low-energy vehicle state. As luck would have it, retargeting actually occurred during the F1 launch and the satellite achieved a satisfactory orbit.

Shortly after being placed into orbit, however, disturbance torques from a nitrogen leak overwhelmed the capacity of the reaction wheels. The satellite spun up in an orientation which prevented the sun from illuminating the solar array, causing an eventual loss in electrical power.

Table 4. Evolution of simulation.

Spacecraft	F1	F2	F3	F4	F5	F6
Spacecraft Type	Block 5D-1					Block 5D-2
Simulation Version Used for IV&V	STF (Software Test Facility)					MSIF
Simulation Location	RCA AFD	CSDL				CSDL & RCA AFD
Host Computer	ID70 (64 kbytes)		PE 8-32C (128 kbytes)			PE 8-32D (512 kbytes)
Flight Computer in Simulation	Single RCA SCP234 (16 kbytes)					Two SCP234/GT (28 kbytes <i>ea</i>)
Simulation Interface Device	CSDL MIH					RCA CIU/GT and GTIU
Launch Booster	MDAC Thor LV-2F with 150 K engine		MDAC Thor LV-2F with 170 K engine configuration			Atlas E configured to the E/F standardized baseline
Significant New Sim Feature	All new	Reentry into Ascent for N ₂ Usage	Orbit Initial Starts	Development of 5D-1 Intermediate Version of MSIF		Redesigned for 5D-2

After approximately three weeks, precession of the F1 spin vector finally permitted USAF communications with the satellite. When experiments indicated that limited control of the spin direction could be accomplished, a DMSP campaign of recovery to operational configuration ensued. The resources for this permitted an STF to be established at CSDL's home facility. This later proved crucial when its ready availability permitted an emergency CSDL redesign of the F1 control laws to a roll-gyroless configuration.

Eventually, a number of flight load reconfigurations were accomplished by the USAF/contractor team, which resulted in the recovery of the vehicle to operational use, including a series of modifications of the control laws to work around IMU gyro failures. Table 5 indicates the major software reconfiguration activities, including episodes on later satellites.

EVOLUTION OF THE CSDL IV&V PROCESS

The IV&V activity evolved from F1 to F6 as a consequence of Block 5D program history and the DMSP community perception of program needs, tempered by the practical realities of limited time and resources. The evolution is summarized in Tables 6 and 7 which show, by comparison against the F6 effort, the effort expended on verification and validation prior to each launch.

Table 5. Block 5D post-launch software reconfigurations.

	Spacecraft Phenomena	Uplinked Software Reconfiguration
F1	Although a satisfactory orbit was achieved during launch, N ₂ control gas leakage spun satellite up to more than 3 rev/min. Satellite electrical power was lost for 3 weeks until spin vector precessed sufficiently for solar arrays to be illuminated	A specialized DESAT program was designed and uplinked to reduce angular momentum and negate average gravity gradient torques using commutated magnetic torquing controlled by earth horizon sensor outputs Recovery programs were created to damp nutation of satellite body and despin satellite to a rate which would not saturate the gyros and be within the control capability of the reaction wheels
	The K7G gyros were damaged by temperature extremes during unpowered spinning. Large drift rates were noted, with roll gyro drift trending toward imminent saturation	An emergency roll-gyroless modification was designed and uplinked to replace the roll gyro control system damping signal with one derived from earth sensor data
	Further deterioration of gyros noted in buildup of drift rates	Flight load reconfigured to implement MOBACS, a pitch momentum bias control law not requiring gyro data, with full dependency on earth sensor outputs for pitch and roll control. Modifications were also made to the magnetic wheel unloading algorithm.
	Frequent toggling noted in sign of magnetic torquers while in magnetic unloading regions	An adjustment was made to the magnetic unloading algorithm to reduce cycling of the roll/yaw magnetic coil
F2	Computer interactions during early mission command operations caused N ₂ thrusting which spun up the spacecraft in orbit	The specialized recovery programs developed for F1 were uplinked and used to despin the spacecraft to its operational configuration
	Excessive, false star mapper outputs due to heavy proton bombardment while over the South Atlantic	Starfix program included to blank out CSA data while over the South Atlantic anomaly region. At same time, an opportunity was taken to adjust control system gains
	Intermittent failure to zero of yaw gyro digitizer	Gyro reasonableness testing incorporated to detect prolonged zero output of yaw gyro and switch to use of redundant skew gyro for yaw control
	Occasional erratic output of sun sensor outputs due to glint from appendages	Reasonableness testing of SSU outputs incorporated to reject spurious data and change mode of operation when yaw error is excessive
F3	Gradual buildup of roll gyro drift toward saturation	A two-gyro mode of operation was uplinked which used skew gyro for roll unless the yaw gyro output was considered unreasonable. If unreasonable, the skew gyro is automatically used for yaw, and an Earth sensor-derived roll rate was used for roll control
	Degradation of earth sensor detector quadrants in the form of increased segment output biases	The ESRS program was modified to provide for individual correction terms for each of the 3 detector segments within each quadrant
F6	Desire for capability of using N ₂ to automatically unload reaction wheel angular momentum. Also, additional reasonableness testing of gyros	Momentum unloading software designed for use on F4 was uplinked
	Desire to eliminate ascent guidance software from orbital flight load to prevent its inadvertent use	After successful launch into orbit, an 8-kbyte block of core containing AGS was zeroed out

Table 6. Flight code verification and validation effort prior to launches.

Flight	Started	Finished	Validation (%)	Verification (%)
F1	Jan. 76	Sept. 76	27	0
F2	Apr. 77	June 77	13	0
F3	Dec. 77	Mar. 78	19	23
F4	June 78	June 79	31	46
F5	Oct. 79	Apr. 80	18	31
F6	June 80	Aug. 82	100	100

Table 7. Evolution of CSDL Block 5D IV&V effort.

Prior to Flight	Flight Code Verification	Flight Code Validation
F1 (1st 5D-1)	Several simulation verification runs, but no formal study of flight code or its data base	Concentration on performance tests, with emphasis on ascent software. Limited orbit software testing. Limited computer plotting capability
F2		Fewer ascent tests. Emphasis on testing of new orbital N ₂ dumping (handback) software
F3	Formal, detailed study of flight code vs. specifications initiated. Considerable increase in simulation verification runs to test failure paths of flight code	Increased emphasis on orbit performance tests using abnormal scenarios. Exhaustive testing of handback software. Instituted computer plotting of all telemetry data
F4	Continuation of detailed study of flight code vs. specifications. A further increase in simulation verification runs	Further emphasis on orbit performance tests using abnormal scenarios and simulated equipment failures. Exhaustive testing of new software (MULS) to employ N ₂ in orbit to unload angular momentum of reaction wheels
F5	Detailed study of 5D-1 flight code, but not data base, was completed	Limited testing of both ascent and orbital software
F6 (1st 5D-2)	Comprehensive study of 5D-2 flight code and data base vs. current specifications was accomplished. Increased emphasis placed on configuration tracking and automatic computer file comparisons	Considerable increase in testing of both ascent and orbit flight code. All paths of mission-critical software were tested. Some de-emphasis of simulation performance testing in favor of verification testing

The F1 CSDL testing experience prior to launch had demonstrated that the ascent flight software would operate properly in a wide range of situations. While the orbit testing was limited to reasonably nominal scenarios, sufficient confidence in the operability of the coding was amassed to permit a launch. During the F1 spin down recovery testing by CSDL, the orbit coding was subjected to a spectrum of anomalous conditions adding to the confidence in both the flight code and the fidelity of the STF.

For F2, minimum changes from F1 in the ascent software resulted in a reduction in the number of ascent runs. The limited time (several months) available for testing was prioritized for focus on new "handback" flight software. Handback was intended to permit the flight IMU and CIU hardware to be returned from orbit to ascent mode to access the N₂ thrusters to either dump residual N₂ or to despin the satellite if a leak should occur.

Unfortunately, the single SCP234 configuration of the STF did not expose a dual flight computer interaction condition which existed. After the launch of F2, the new handback feature was used in an attempt to latch, by N₂ thrusting, the solar array which had failed to complete its deployment sequence. The computer controlling the handback process commanded the flight equipment (IMU and CIU) to ascent mode. However, the other computer interfered by commanding the equipment back to orbit mode in which a different IMU data rate and scaling was used. As a result, the satellite was spun up by the handback software. The vehicle was subsequently restored to an operational state using the special software loads which had already been developed for the recovery of F1. This incident increased program awareness of the hazards of the computer interactions, and resources were provided for CSDL to start development of the dual computer MSTF.

During the launch of F2, the IMU was subjected to unusually high (4 sigma), short duration vibrational disturbances at liftoff, causing software overflows in the ascent guidance software (AGS). The consequence was an improper orbit inclination, affecting the precessional rate of the orbit about the Earth's polar axis. As a result, the AGS coding for F3 was rescaled to accommodate IMU saturation conditions, and the CSDL effort was expanded to increase the level of testing and to initiate a formal campaign of detailed flight code analysis.

The flight code analysis involved a line-by-line walk-through of program assembly listings and comparison against the logic shown in available specifications. Discrepancies between code and specifications were reported, but up-to-date specifications for some software modules were not always existent for meaningful comparisons.

For F3 testing, additional time and resources were made available so that the standard repertoire of ascent and orbit tests was augmented by runs in which special scenarios exercised many of the off-nominal paths in the orbit code. For F3, automatic post-run plotting of all telemetry data was implemented to facilitate analysis of the enormous amount of data from each run. Up to that time, analysts had to wade through piles of formatted dumps of telemetry data, hand plotting selective quantities of interest. The launch and performance of F3 were superbly uneventful.

With several useful satellites now in orbit, time existed for more comprehensive STF testing by CSDL of the F4 flight code, so a larger number of the off-nominal paths were exercised by performance runs in a variety of scenarios (such as failures of gyros, earth sensor quadrants, reaction wheels, momentum unloading coils in both normal and abnormal satellite attitudes). For F4, the N_2 hand-back code was redesigned by RCA to enable the use of N_2 for automatic unloading of reaction wheel angular momentum (in parallel with the capability already provided by the MMU coils). This new feature, called MULS, was exhaustively tested by CSDL.

The F4 version of flight code was essentially frozen for F5 so that the CSDL F5 series of tests was limited to the standard ascent series of tests (because of the criticality of the launch portion of the mission) and several nominal orbit tests.

Development of MSTF, a dual SCP simulation for 5D-2, was completed and ready for use by CSDL for the F6 V&V activity with a new generation of flight software. It provided a number of new tools to enhance the conduct of tests and analysis of data. (See Table 3.)

The IV&V methodology employed by CSDL for F6 was summarized in Tables 1 and 2. This methodology included considerable additional simulation testing, V&V of the flight code data base, routine file dump comparisons on a module-by-module basis of all flight load deliveries, and semiautomatic specification and code configuration tracking. For F6, sufficient time, resources, and documentation were available to verify the flight code against meaningful design specifications, validate all flight code functions by simulation, and exercise all paths in mission-critical software modules (such as ascent guidance software and new N_2 momentum unloading software, GULS) prior to launch. The launch was, of course, successful and satellite performance conformed to that predicted by the MSTF simulation.

CONCLUSIONS

The Block 5D IV&V effort totally paid for itself in the launch of the first 5D-1 vehicle. In subsequent launches and a variety of software reconfigurations, it uncovered a number of software problems and generally provided comfort to the DMSP community. Additional confidence was provided as required by the program atmosphere by increasing the level of IV&V.

During the conduct of the CSDL Block 5D IV&V work and simulation development effort which accompanied it, a number of lessons were learned by the CSDL team which may be helpful for similar work in the future.

First, regarding software validation and verification:

- (1) IV&V requires intimate knowledge of the software and system. To develop this expertise and to ensure that effective IV&V is even possible, the activity should be initiated as early as possible in the program development cycle.
- (2) Software IV&V is a systems task requiring systems knowledge and experience, not just software expertise.
- (3) There is no substitute for the validation technique of operating the actual software in a closed-loop simulation. This provides the highest overall confidence in the usability of the software.
- (4) Software validation is not in itself sufficient; verification is also required to find software problems.
- (5) Every opportunity should be exploited to obtain machine-readable information and to insert that data into a computer for it can be sorted and manipulated in a remarkable number of ways.
- (6) Frequently, the most interesting results of a test will be unrelated to the test objective.
- (7) Some iteration of the IV&V effort can always be expected due to either design changes or corrections to problems found during V&V.
- (8) Simulation tools can be used in a variety of unanticipated ways provided that ready access to these facilities is available.

There are also some observations regarding simulation development that may be useful for future efforts.

- (1) The documentation of a computer program is as important as the coding itself, for the program will surely be redesigned and used by someone other than the original designer and programmer.
- (2) More computer memory is always needed, and getting the most possible at the start will result in straightforward coding and lower overall software development costs.
- (3) While the current trend is for virtually exclusive use of higher-order languages to minimize development cost, assembly languages for simulations provide the benefit of facilitating patching for special purposes.
- (4) Software configuration control is indispensable for preserving design integrity.
- (5) While there is a trend toward elimination of traditional software flow diagrams in favor of documentation by program development languages, flow diagrams are indispensable in involving nonprogrammer personnel in the design and review process.

- (6) Design assumptions and references should be carefully documented because they will surely be challenged, reviewed, and re-reviewed at some future time.
- (7) Extra fidelity in modeling will usually have a payoff.

REFERENCES

1. USAF Fact Sheet 82-58, Defense Meteorological Satellite Program.
2. USAF SAMSO pamphlet 800-4, "Acquisition Management, Independent Stability and Control Analysis of Orbit, Reentry, and Launch Vehicles," 15 Feb. 1975.
3. Beck, George A., "Precision Spacecraft Attitude Control Using an On-Board Computer," EASCON '75, Sept. 29-Oct. 1, 1975.
4. Beck, Buntschuh, and Aukstikalnis, "Adaptable Spacecraft Computer Systems Experience," IFAC Conference, Cambridge, Mass., 9 Sept. 1980.
5. Hursh, John W., "Flight Software Validation Through Interactive Closed-Loop System Simulation," 13th International Symposium on Space Technology and Science, Tokyo, 1982.
6. Mayfield, J., "Weather Satellite Production Continues," *Aviation Week and Space Technology*, 4 June 1979, pp. 47-58.
7. "New Weather Satellite Series Planned," *Aviation Week and Space Technology*, 15 July 1974, pp. 41-47.
8. Smith, Msgt Gerald H., "Block 5D Was Born Again," *Airman*, Jan. 1978, pp. 45-48.
9. Staniszewski, J. R., "Lazarus Sleeps No More - Satellite Recovered From Tumble," *RCA Engineer*, Aug./Sept. 1978, pp. 35-43.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document						
	AGARD-CP-350	ISBN 92-835-0349-X	UNCLASSIFIED						
5. Originator	Advisory Group for Aerospace Research and Development North Atlantic Treaty Organization 7 rue Ancelle, 92200 Neuilly sur Seine, France								
6. Title	GUIDANCE AND CONTROL TECHNIQUES FOR ADVANCED SPACE VEHICLES								
7. Presented at	the Guidance and Control Panel 37th Symposium held in Florence, Italy, 27-30 September 1983								
8. Author(s)/Editor(s)	9. Date								
Various	January 1984								
10. Author's/Editor's Address	11. Pages								
Various	356								
12. Distribution Statement	This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.								
13. Keywords/Descriptors	<table border="0"> <tr> <td>Spacecraft control</td> <td>Flexible spacecraft</td> </tr> <tr> <td>Docking rendezvous</td> <td>Mission control systems</td> </tr> <tr> <td>Space structures</td> <td>Attitude control</td> </tr> </table>			Spacecraft control	Flexible spacecraft	Docking rendezvous	Mission control systems	Space structures	Attitude control
Spacecraft control	Flexible spacecraft								
Docking rendezvous	Mission control systems								
Space structures	Attitude control								
14. Abstract	<p>✓ This publication contains most of the papers presented at the 37th Guidance and Control Panel Symposium held in Florence, Italy on 27-30 September 1983.</p> <p>27 papers were programmed, two of which are printed in GCP/CP-350 Supplement, classified NATO-CONFIDENTIAL.</p> <p>The papers covered the following topics; Review of mission requirements and technology issues; Control components; Estimation and control, Orbital manoeuvring; Large space structures; Test and performance evaluation.</p>								

<p>AGARD Conference Proceedings No.350 Advisory Group for Aerospace Research and Development, NATO GUIDANCE AND CONTROL TECHNIQUES FOR ADVANCED SPACE VEHICLES Published January 1984 356 pages</p> <p>This publication contains most of the papers presented at the 37th Guidance and Control Panel Symposium held in Florence, Italy on 27-30 September 1983.</p> <p>27 papers were programmed, two of which are printed in GCP/CP-350 Supplement, classified NATO-CONFIDENTIAL.</p> <p>P.T.O.</p>	<p>AGARD-CP-350</p> <p>Spacecraft control Docking rendezvous Space structures Flexible spacecraft Mission control systems Attitude control</p>	<p>AGARD Conference Proceedings No.350 Advisory Group for Aerospace Research and Development, NATO GUIDANCE AND CONTROL TECHNIQUES FOR ADVANCED SPACE VEHICLES Published January 1984 356 pages</p> <p>This publication contains most of the papers presented at the 37th Guidance and Control Panel Symposium held in Florence, Italy on 27-30 September 1983.</p> <p>27 papers were programmed, two of which are printed in GCP/CP-350 Supplement, classified NATO-CONFIDENTIAL.</p> <p>P.T.O.</p>	<p>AGARD-CP-350</p> <p>Spacecraft control Docking rendezvous Space structures Flexible spacecraft Mission control systems Attitude control</p>
<p>AGARD Conference Proceedings No.350 Advisory Group for Aerospace Research and Development, NATO GUIDANCE AND CONTROL TECHNIQUES FOR ADVANCED SPACE VEHICLES Published January 1984 356 pages</p> <p>This publication contains most of the papers presented at the 37th Guidance and Control Panel Symposium held in Florence, Italy on 27-30 September 1983.</p> <p>27 papers were programmed, two of which are printed in GCP/CP-350 Supplement, classified NATO-CONFIDENTIAL.</p> <p>P.T.O.</p>	<p>AGARD-CP-350</p> <p>Spacecraft control Docking rendezvous Space structures Flexible spacecraft Mission control systems Attitude control</p>	<p>AGARD Conference Proceedings No.350 Advisory Group for Aerospace Research and Development, NATO GUIDANCE AND CONTROL TECHNIQUES FOR ADVANCED SPACE VEHICLES Published January 1984 356 pages</p> <p>This publication contains most of the papers presented at the 37th Guidance and Control Panel Symposium held in Florence, Italy on 27-30 September 1983.</p> <p>27 papers were programmed, two of which are printed in GCP/CP-350 Supplement, classified NATO-CONFIDENTIAL.</p> <p>P.T.O.</p>	<p>AGARD-CP-350</p> <p>Spacecraft control Docking rendezvous Space structures Flexible spacecraft Mission control systems Attitude control</p>

<p>The papers covered the following topics: Review of mission requirements and technology issues; Control components; Estimation and control, Orbital manoeuvring; Large space structures; Test and performance evaluation.</p>	<p>The papers covered the following topics: Review of mission requirements and technology issues; Control components; Estimation and control, Orbital manoeuvring; Large space structures; Test and performance evaluation.</p>
<p>ISBN 92-835-0349-X</p> <p>The papers covered the following topics: Review of mission requirements and technology issues; Control components; Estimation and control, Orbital manoeuvring; Large space structures; Test and performance evaluation.</p>	<p>ISBN 92-835-0349-X</p> <p>The papers covered the following topics: Review of mission requirements and technology issues; Control components; Estimation and control, Orbital manoeuvring; Large space structures; Test and performance evaluation.</p>
<p>ISBN 92-835-0349-X</p>	<p>ISBN 92-835-0349-X</p>

END

FILMED

7 - 81

DTIC

AD-A141 969

CONFERENCE PROCEEDINGS ON GUIDANCE AND CONTROL
TECHNIQUES FOR ADVANCED SP..(U) ADVISORY GROUP FOR
AEROSPACE RESEARCH AND DEVELOPMENT NEUILLY.. JAN 84
AGARD-CP-350

5/5

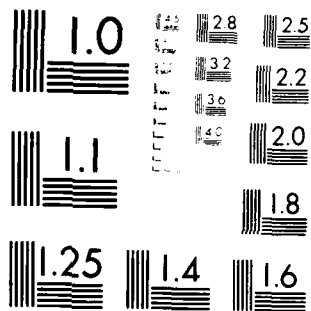
UNCLASSIFIED

F/G 22/3

NL



END
DATE
FILMED
9-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SUPPLEMENTARY

INFORMATION



DEPARTMENT OF THE AIR FORCE
HEADQUARTERS UNITED STATES AIR FORCE
WASHINGTON, D C 20330

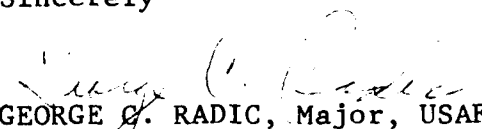
28 JUN 1984

AD-A141969
Mr. J. E. Cundiff
Defense Technical Information Center
Defense Logistics Agency
Cameron Station
Alexandria, VA 22314

Dear Mr. Cundiff

Regarding AGARD-CP-350, Paper No. 16, by Bowles and Altobelli, the proprietary information statement can be removed. Confirmation is provided in the attached AGARD letter, which includes the Clearance Certificate.

Sincerely


GEORGE C. RADIC, Major, USAF
US National Coordinator for
AGARD

1 Atch
NATO AGARD Ltr, 28 May 84
w/Atch

File



AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT
NORTH ATLANTIC TREATY ORGANIZATION

ETS/HL/92

28 May 1984

Major G C Radic
Headquarters United States Air Force
Attn: AF/RDI
The Pentagon
Washington DC 20330

Dear Major Radic

I write in reply to your letter of 9 May.

Regarding CP-350, Paper No 16, by Bowles and Altobelli, I confirm that the proprietary information statement should have been removed prior to printing. As you will see from the attachment, we hold a Clearance Certificate which states specifically that the Paper contains no proprietary information. You may therefore instruct DTIC to remove the statement.

Concerning the AGARD Index 80-82, we never know the precise dates of hard-copy shipments to the US. We merely know that the cartons are delivered to the Air Base*. It is a matter of speculation just how long they have to stay there before they are transported since shipments are made on a 'space-available' basis. However, the facts are as follows: the bulk hard-copy consignment for NASA Langley was delivered to the Air Base on 20 February 1984. All other US copies (that is, for the National Delegates, Panel Members, etc.) were sent by the APO on 18 February. No copies were sent to NTIS because we were advised last year (see letter attached) that NTIS no longer requires copies of AGARD publications to be provided direct.

The microfiche copies of the NATO Index were all dispatched to STIF, Baltimore/Washington International Airport, on 16 May by normal post.

Yours sincerely

E.T. Sharp

E T Sharp
Scientific Publications
Executive

Attachments:
- As stated

*Mildenhall, in the UK

PAPER CLEARANCE REQUEST

1. AUTHOR: Bowles, W. M.	2. PROJECT NO: 40-91-00	3. DDC NO: S33 / C377	4. ALHORN SIGNATURE: <i>W. M. Bowles</i>	5. DATE: 15 August 1983
Altobelli, M. R.	40-92-10	S33 / C377	<i>Mark Altobelli</i>	
		1		

6. TITLE (MUST BE UNCLASSIFIED): **AN APPROACH TO ATTITUDE CONTROL WHICH HELPS MAINTAIN ORBIT INCLINATION**

7. ABSTRACT ☐

8. COMPLETE PAPER ☒

9. PURPOSE: **8. SPECIFIC INFORMATION, MEET, SPONSOR DATE, LOCATION, JOURNAL, PUBLICATION**

9. DEADLINE FOR SUBMITTAL TO PROGRAM CHAIRMAN: **15 August 1983**

PRESENTATION ☐ **NATO, (AGARD) CONFERENCE ON SATELLITE ATTITUDE CONTROL**

PUBLICATION ☐ **September 27-30, 1983**

BOTH ☒ **Florence, Italy**

11. PROPRIETARY INFORMATION YES ☐ NO ☒ 12. PATENTABLE INFORMATION YES ☐ NO ☒ 13. RELATED PATENT DISCLOSURE NO'S.

COMPLETE FOLLOWING IF INFORMATION CONTRACT CONNECTED: CHECK CURRENT "CONTRACT SECURITY CLASSIFICATION SPECIFICATION" FOR SECURITY CLASSIFICATION.

14. CONTRACT NUMBER: INTEL-153	15. PRODUCING AGENCY, LOCATION: INTELSAT	16. COGNIZANT GOVT PROJ MONITOR: YES	17. CODE: YES	18. DD FORM 124: YES
CONTRACT NUMBER	PRODUCING AGENCY, LOCATION	COGNIZANT GOVT PROJ MONITOR	CODE	DD FORM 254: YES

COMPLETE THE FOLLOWING IF THIS PAPER ☐ OR AND ABSTRACT OF IT ☐ (CHECK ONE), HAS BEEN PREVIOUSLY CLEARED:

DATE: _____ TITLE AND AUTHOR: _____

REF NO: **501015**

REMARKS:

19. DOCUMENT REVIEWED AND APPROVED FOR RELEASE	20. IF CONTRACT RELATED	21. IF RELATED TO CURRENT OR PLANNED PRODUCT LINE
<i>W. M. Bowles</i> 8/9/83	<i>W. M. Bowles</i> 8/9/83	<i>W. M. Bowles</i> 8/9/83
IMMEDIATE SUPP. DATE	PROJECT MON. DATE	PRODUCT LINE DATE
<i>W. M. Bowles</i> 8/9/83	<i>W. M. Bowles</i> 8-9-83	<i>W. M. Bowles</i> 8/9/83
ATTACHMENT NO. DATE	PROGRAM MON. DATE	FINAL, SOON TO BE FINAL DATE
<i>W. M. Bowles</i> 8/9/83	<i>W. M. Bowles</i> 8-9-83	<i>W. M. Bowles</i> 8/9/83

22. TO BE COMPLETED BY PROGRAM MON. OR ORIGINATOR:

<input checked="" type="checkbox"/> HUGHES INTERNATIONAL	<i>W. M. Bowles</i> 19 Aug 83
<input type="checkbox"/> NOT APPLICABLE TO HIM	APPROVED (PC SIGNS IF NOT APPLICABLE)
<input checked="" type="checkbox"/> INTELSAT	<i>A. E. LELVEN</i> 12 Aug 83
GOVERNMENT	PC SIGNS AND STATES SEPARATE REJECTION OR APPROVAL
<input type="checkbox"/> GOVERNMENT	PC SIGNS AND STATES SEPARATE REJECTION OR APPROVAL
<input checked="" type="checkbox"/> PATENTS-CORPORATE OFFICE	<i>William F. ...</i> 8-22-83
<input checked="" type="checkbox"/> PUBLIC RELATIONS & ADVERTISING - CORPORATE OFFICE	<i>J. H. ...</i> 8-22-83
<input checked="" type="checkbox"/> COMMENTS (LIMITATIONS OR REASONS FOR DENIAL)	<i>Bowles submission of disclosure.</i>

ACTION COMPLETED ☒ RELEASE APPROVED ☐ DENIED RELEASE

W. M. Bowles 22 Aug 1983

DISTRIBUTION: WHITE-PUBLICATION CLEARANCE OFFICE, CANARY & PINK-PATENT CORP. OFFICE, BLUE-PUBLIC RELATIONS CORP. OFFICE, GREEN-PUBLICATION CLEARANCE FILE, GOLDENROD-ORIGINATOR/AUTHOR

IDENTIFICATION

NAME

ADDRESS

CITY

STATE

DATE OF BIRTH

AT APPROACH TO ATTITUDE OF THE

PRESENTATION

PUBLICATION

NOTE

PERIODICAL

INFORMATION

COMPLETE FOLLOW-UP

CONTROL

INTER-DEPT

CONTRACT

COMPLETE THE FOLLOWING

DATE

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

REMARKS

DATE
ILME